

Search completed: November 10, 2004, 13:38:22
Job time : 43.2953 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 12:32:37 ; Search time 10.7081 Seconds
(without alignments)
191.991 Million cell updates/sec

Title: US-10-092-750-26

Perfect score: 148

Sequence: 1 ANLLLMVPIILAMFLMTERKILGYIQPR 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA:**
1: /cgn2_6/ptodata/1/iaa/5A COMB.pap.*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pap.*
3: /cgn2_6/ptodata/1/iaa/5C COMB.pap.*
4: /cgn2_6/ptodata/1/iaa/5D COMB.pap.*
5: /cgn2_6/ptodata/1/iaa/5E COMB.pap.*
6: /cgn2_6/ptodata/1/iaa/5F COMB.pap.*
7: /cgn2_6/ptodata/1/iaa/5G COMB.pap.*
8: /cgn2_6/ptodata/1/iaa/5H COMB.pap.*
9: /cgn2_6/ptodata/1/iaa/5I COMB.pap.*
10: /cgn2_6/ptodata/1/iaa/5J COMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	134	90.5	83	1	US-08-224-983-2
2	134	90.5	83	2	US-08-852-933-2
3	134	90.5	83	2	US-08-852-945-2
4	134	90.5	83	2	US-08-852-945-2
5	134	90.5	83	3	US-08-852-865-2
6	131	88.5	318	3	US-09-097-889-13
7	131	88.5	318	4	US-09-098-079-13
8	87	58.8	316	4	US-09-270-767-41860
9	83	56.1	286	4	US-09-270-767-44872
10	80	54.1	110	4	US-09-248-796A-14320
11	63	42.6	371	2	US-08-928-692-20
12	63	42.6	371	3	US-09-339-972-20
13	53	35.8	189	4	US-09-270-767-41238
14	53	35.8	189	4	US-09-270-767-56454
15	52	35.1	850	4	US-09-583-110-4394
16	50	33.8	513	3	US-09-097-889-15
17	50	33.8	513	4	US-09-098-079-15
18	47	31.8	367	4	US-09-543-681A-4184
19	46	31.1	178	4	US-09-489-039A-11561
20	46	31.1	210	4	US-09-248-796A-25835
21	46	31.1	361	1	US-08-415-751-7
22	46	31.1	372	4	US-09-107-532A-4779
23	46	31.1	413	4	US-09-328-352-5589
24	46	31.1	686	2	US-08-993-228-12
25	46	31.1	2910	1	US-08-466-033-183
26	46	31.1	2910	2	US-08-444-733-183
27	46	31.1	2910	2	US-08-464-134-183

28	46	31.1	2910	2	US-08-461-361-183	Sequence 183, App
29	46	31.1	2910	2	US-08-485-910-183	Sequence 183, App
30	46	31.1	2910	5	PCT-US95-06266-157	Sequence 157, App
31	45.5	30.7	339	4	US-09-328-352-4674	Sequence 4674, App
32	45.5	30.7	459	4	US-09-129-112-9	Sequence 9, Appli
33	45	30.4	161	4	US-09-248-796A-15595	Sequence 15595, A
34	45	30.4	320	4	US-09-248-796A-14837	Sequence 14837, A
35	45	30.4	332	4	US-09-489-039A-12580	Sequence 12580, A
36	45	30.4	364	4	US-09-252-991A-31716	Sequence 31716, A
37	45	30.4	418	4	US-09-252-991A-17598	Sequence 17598, A
38	45	30.4	423	4	US-09-540-236-3149	Sequence 3149, App
39	45	30.4	442	4	US-09-328-352-5748	Sequence 5748, App
40	45	30.4	510	4	US-09-248-796A-20738	Sequence 20738, A
41	45	30.4	515	4	US-09-170-496D-104	Sequence 104, App
42	45	30.4	515	4	US-09-170-496D-220	Sequence 220, App
43	45	30.4	520	4	US-09-248-796A-20803	Sequence 20803, A
44	44.5	30.1	124	4	US-09-248-796A-15423	Sequence 15423, A
45	44	29.7	114	4	US-09-710-279-2298	Sequence 2298, App

ALIGNMENTS

RESULT 1
US-08-224-983-2
; Sequence 2, Application US/08224983
; Patent No. 5646011
; GENERAL INFORMATION:
; APPLICANT: Yokoyama, Shiro
; TITLE OF INVENTION: Cisplatin Resistance Gene and Uses Therefor
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/224,983
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Giulio A. DeConti, Jr.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: BBI-010
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 83 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-224-983-2

Query Match 90.5%; Score 134; DB 1; Length 83;
Best Local Similarity 93.5%; Pred. No. 2.6e-13;
Matches 29; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ANLLLMVPIILAMFLMTERKILGYIQPR 31
|||
DB 2 ANLLLMVPIILAMFLMTERKILGYIQPR 32
|||

RESULT 2
US-08-852-933-2

```

; Sequence 2, Application US/08852933
; Patent No. 5846725
; GENERAL INFORMATION:
; APPLICANT: Yokoyama, Shiro
; TITLE OF INVENTION: Cisplatin Resistance Gene and Uses Therefor
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, suite 510
; City: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/852,933
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/224,983
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Giulio A. DeConti, Jr.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: BBI-010
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 83 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-852-933-2

Query Match 90.5%; Score 134; DB 2; Length 83;
Best Local Similarity 93.5%; Pred. No. 2.6e-13;
Matches 29; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ANLLLMVPIILIAMAFMLTKRKILGYIQPR 31
Db 2 ANLLLIIVPIILIAMAFMLTKRKILGYIQLR 32

RESULT 3
US-08-852-945-2
; Sequence 2, Application US/08852945
; Patent No. 5858674
; GENERAL INFORMATION:
; APPLICANT: Yokoyama, Shiro
; TITLE OF INVENTION: Cisplatin Resistance Gene and Uses Therefor
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, suite 510
; City: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/852,945
; FILING DATE: 08-MAY-1997
; CLASSIFICATION: 435

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/224,983
; FILING DATE: 08-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Giulio A. DeConti, Jr.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: BBI-010
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 83 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-852-945-2

Query Match 90.5%; Score 134; DB 2; Length 83;
Best Local Similarity 93.5%; Pred. No. 2.6e-13;
Matches 29; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ANLLLMVPIILIAMAFMLTKRKILGYIQPR 31
Db 2 ANLLLIIVPIILIAMAFMLTKRKILGYIQLR 32

RESULT 4
US-08-853-021-2
; Sequence 2, Application US/08853021
; Patent No. 5858780
; GENERAL INFORMATION:
; APPLICANT: Yokoyama, Shiro
; TITLE OF INVENTION: Cisplatin Resistance Gene and Uses Therefor
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, suite 510
; City: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/853,021
; FILING DATE: 08-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/224,983
; FILING DATE: 08-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Giulio A. DeConti, Jr.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: BBI-010
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 83 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-853-021-2

Query Match 90.5%; Score 134; DB 2; Length 83;
Best Local Similarity 93.5%; Pred. No. 2.6e-13;
Matches 29; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```


COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/098,079
FILING DATE: 15-JUN-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Rosenman, Ph.D., Stephen J.
REGISTRATION NUMBER: 43,058
REFERENCE/DOCKET NUMBER: 660088.416
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:

```
; LENGTH: 318 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-09-099-079-13

Query Match      88.5%; Score 131; DB 4; Length 318;
Best Local Similarity 90.3%; Pred. No. 3.4e-12;
Matches 28; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ANLLLMVPIIIAFAFLMLTERKILGYIQPR 31
   |||||:|||||:|||||:|||||:|||||:
DB 4 ANLLLIIVPIIIAFAFLMLTERKILGYMQLR 34
   |||||:|||||:|||||:|||||:|||||:

RESULT 8
US-09-270-767-41860
; Sequence 41860, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 2.0
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41860
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-41860

Query Match      58.8%; Score 87; DB 4; Length 316;
Best Local Similarity 60.7%; Pred. No. 1.8e-05;
Matches 17; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

QY 4 LLLMVPIIIAFAFLMLTERKILGYIQPR 31
   |||||:|||||:|||||:|||||:|||||:
DB 15 LLLIICVLVSVAFLTLRKKVIGYIQIR 42
   |||||:|||||:|||||:|||||:|||||:

RESULT 9
US-09-270-767-44872
; Sequence 44872, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 44872
; LENGTH: 286
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-44872

Query Match      56.1%; Score 83; DB 4; Length 286;
Best Local Similarity 57.1%; Pred. No. 6.7e-05;
Matches 16; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

QY 4 LLLMVPIIIAFAFLMLTERKILGYIQPR 31
   |||||:|||||:|||||:|||||:|||||:
DB 22 LLLIICVLVSVAFLTLRKKVIGYIQIR 49
   |||||:|||||:|||||:|||||:|||||:
```

US-08-928-692-20

Query Match 42.6%; Score 63; DB 2; Length 371;
Best Local Similarity 41.4%; Pred. No. 0.1;
Matches 12; Conservative 10; Mismatches 7; Indels 0; Gaps 0;

QY 3 LLLVLPILIAMAFMLTERKILGYIOPR 31
DB 13 VLLVLPSSLLAVAVTVVAERTKMTASMQRR 41

RESULT 12

US-09-339-972-20
; Sequence 20, Application US/09339972
; Patent No. 6323002

GENERAL INFORMATION:

APPLICANT: Brody, Howard

APPLICANT: Yaver, Deborah S.

APPLICANT: Lamsa, Michael

APPLICANT: Hansen, Kim

TITLE OF INVENTION: Methods for Modifying the Production of

TITLE OF INVENTION: a Polypeptide

NUMBER OF SEQUENCES: 80

CORRESPONDENCE ADDRESS:

ADDRESSEE: No. 63230020 No. 6323002disk of No. 6323002th America, Inc.

STREET: 405 Lexington Avenue

CITY: New York

STATE: NY

COUNTRY: USA

ZIP: 10174

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/339,972

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/928,692

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Lambiris, Elias J

REGISTRATION NUMBER: 33,728

REFERENCE/DOCKET NUMBER: 4944.200-US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-867-0123

TELEFAX: 212-878-9655

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:

LENGTH: 371 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: No. 6323002e

US-09-339-972-20

Query Match 42.6%; Score 63; DB 3; Length 371;
Best Local Similarity 41.4%; Pred. No. 0.1;
Matches 12; Conservative 10; Mismatches 7; Indels 0; Gaps 0;

QY 3 LLLVLPILIAMAFMLTERKILGYIOPR 31
DB 13 VLLVLPSSLLAVAVTVVAERTKMTASMQRR 41

RESULT 13

US-09-270-767-41238

; Sequence 41238, Application US/09270767

; Patent No. 6703491

; GENERAL INFORMATION:

; APPLICANT: Homburger et al.

; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41238
; LENGTH: 189
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-41238

Query Match 35.8%; Score 53; DB 4; Length 189;
Best Local Similarity 43.5%; Pred. No. 1.6;
Matches 10; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 ANLLLMVPILIAMAFMLTERK 23
DB 77 ANVLLRLKFLAMAFIVAKQK 99

RESULT 14

US-09-270-767-56454

; Sequence 56454, Application US/09270767

; Patent No. 6703491

; GENERAL INFORMATION:

; APPLICANT: Homburger et al.

; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

; FILE REFERENCE: File Reference: 7328-094

; CURRENT APPLICATION NUMBER: US/09/270,767

; CURRENT FILING DATE: 1999-03-17

; NUMBER OF SEQ ID NOS: 62517

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 56454

; LENGTH: 189

; TYPE: PRT

; ORGANISM: Drosophila melanogaster

; FEATURE:

; OTHER INFORMATION: Xaa means any amino acid

US-09-270-767-56454

Query Match 35.8%; Score 53; DB 4; Length 189;
Best Local Similarity 43.5%; Pred. No. 1.6;
Matches 10; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 ANLLLMVPILIAMAFMLTERK 23
DB 77 ANVLLRLKFLAMAFIVAKQK 99

RESULT 15

US-09-583-110-4394

; Sequence 4394, Application US/09583110

; Patent No. 6699703

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al.

; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus

; FILE REFERENCE: Pneumoniae for Diagnostics and Therapeutics

; CURRENT APPLICATION NUMBER: US/09/583,110

; CURRENT FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: US 09/107,433

; PRIOR FILING DATE: 1998-06-30

; PRIOR APPLICATION NUMBER: US 60/085,131

; PRIOR FILING DATE: 1998-05-12

; PRIOR APPLICATION NUMBER: US 60/051,553

; PRIOR FILING DATE: 1997-07-02

; NUMBER OF SEQ ID NOS: 5322

; SEQ ID NO 4394

; LENGTH: 850

; TYPE: PRT

```

; ORGANISM: Streptococcus pneumoniae
US-09-583-110-4394
Query Match      35.1%; Score 52; DB 4; Length 850;
Best Local Similarity 33.3%; Pred. No. 13;
Matches 8; Conservative 10; Mismatches 6; Indels 0; Gaps 0;

QY      4 LLLMVPILIAAFMLTERKILGY 27
      : : : : : | : : : : : | : : : : :
Db      161 VFILPLIITGLHLITEKLLLY 184

```

Search completed: November 10, 2004, 13:44:01
Job time : 11.7081 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 16:36:12 ; Search time 33.3022 Seconds
(without alignments)
328.807 Million cell updates/sec

Title: US-10-092-750-26

Perfect score: 148
Sequence: 1 ANLLLMVPILIAMFLMTERKILGYIQPR 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225986 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	148	100.0	31 14	US-10-092-750-26
2	131	88.5	318 9	US-09-098-079-13
3	131	88.5	318 15	US-10-428-487-34
4	131	88.5	318 16	US-10-408-765A-189
5	130	87.8	101 15	US-10-264-049-3171
6	130	87.8	261 17	US-10-425-115-310657
7	127	85.8	318 16	US-10-408-765A-1490
8	120	81.1	79 15	US-10-264-049-3416
9	111.5	75.3	110 15	US-10-424-599-226150
10	65	43.9	355 15	US-10-282-122A-47875
11	64	43.2	46 17	US-10-425-115-332406
12	63	42.6	371 17	US-10-000-845-20
13	60	40.5	134 17	US-10-425-115-202752
				Sequence 26, Appl
				Sequence 13, Appl
				Sequence 34, Appl
				Sequence 189, Appl
				Sequence 3171, Ap
				Sequence 310657,
				Sequence 1490, Ap
				Sequence 3416, Ap
				Sequence 226150,
				Sequence 47875, A
				Sequence 332406,
				Sequence 20, Appl
				Sequence 202752,

RESULT 2

US-09-098-079-13
; Sequence 13, Application US/09098079
; Patent No. US20020064773A1
; GENERAL INFORMATION:

Query Match 100.0%; Score 148; DB 14; Length 31;
Best Local Similarity 100.0%; Pred. No. 3.3e-14;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANLLLMVPILIAMFLMTERKILGYIQPR 31

Db 1 ANLLLMVPILIAMFLMTERKILGYIQPR 31

ALIGNMENTS

RESULT 1

US-10-092-750-26
; Sequence 26, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; APPLICANT: Wright, Martin C.
; TITLE OF INVENTION: Polypeptides Interactive with BCL-X1
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092,750
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/274,526
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-26

14	60	40.5	138	15	US-10-424-599-161174	Sequence 161174,
15	55	37.2	261	15	US-10-282-122A-45946	Sequence 45946, A
16	55	37.2	263	15	US-10-282-122A-45895	Sequence 45895, A
17	52	35.1	204	16	US-10-781-014-660	Sequence 660, App
18	52	35.1	552	16	US-10-781-014-654	Sequence 654, App
19	52	35.1	581	9	US-09-738-626-8269	Sequence 6269, Ap
20	52	35.1	584	9	US-09-945-825-2	Sequence 2, Appli
21	52	35.1	584	15	US-10-380-055-2	Sequence 2, Appli
22	52	35.1	850	16	US-10-474-776-421	Sequence 421, App
23	50.5	34.1	454	14	US-10-156-761-12378	Sequence 12378, A
24	50	33.8	44	15	US-10-264-049-3246	Sequence 3246, Ap
25	50	33.8	229	13	US-10-079-623-359	Sequence 359, App
26	50	33.8	513	9	US-09-098-079-15	Sequence 15, Appl
27	50	33.8	513	15	US-10-428-487-32	Sequence 32, Appl
28	50	33.8	513	15	US-10-231-956A-64	Sequence 64, Appl
29	50	33.8	513	16	US-10-408-765A-191	Sequence 191, App
30	50	33.8	513	17	US-10-770-668-72	Sequence 72, Appl
31	49	33.1	16	16	US-10-777-053-799	Sequence 799, App
32	49	33.1	16	17	US-10-837-217-799	Sequence 799, App
33	49	33.1	50	11	US-09-864-408A-4702	Sequence 4702, Ap
34	49	33.1	50	17	US-10-425-115-345823	Sequence 345823, A
35	49	33.1	176	16	US-10-767-701-46417	Sequence 46417, A
36	49	33.1	227	17	US-10-425-115-276127	Sequence 276127,
37	49	33.1	324	16	US-10-437-963-181641	Sequence 181641,
38	49	33.1	489	9	US-09-738-626-3576	Sequence 3576, Ap
39	49	33.1	564	15	US-10-282-122A-54166	Sequence 54166, A
40	49	33.1	624	14	US-10-389-493-16411	Sequence 16411, A
41	48.5	32.8	329	9	US-09-895-913A-170	Sequence 170, App
42	48.5	32.8	329	15	US-10-335-977-5684	Sequence 5684, Ap
43	48.5	32.8	1894	14	US-10-369-493-2252	Sequence 2252, Ap
44	48	32.4	92	16	US-10-767-701-36902	Sequence 36902, A
45	48	32.4	164	15	US-10-424-599-205152	Sequence 205152,

; APPLICANT: Herrnstadt, Corrina
 ; APPLICANT: Ghosh, Soumitra S.
 ; APPLICANT: Clevenger, William
 ; APPLICANT: Fahy, Eoin F.
 ; APPLICANT: Davis, Robert E.
 ; TITLE OF INVENTION: DIAGNOSTIC METHOD BASED ON QUANTIFICATION OF
 ; TITLE OF INVENTION: EXTRAMITOCHONDRIAL DNA
 ; NUMBER OF SEQUENCES: 26
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: SED and BERRY LLP
 ; STREET: 6300 Columbia Center, 701 Fifth Avenue
 ; CITY: Seattle
 ; STATE: Washington
 ; COUNTRY: USA
 ; ZIP: 98104
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/098,079
 ; FILING DATE: 15-JUN-1998
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Roserman Ph.D., Stephen J.
 ; REGISTRATION NUMBER: 43,058
 ; REFERENCE/DOCKET NUMBER: 660088.416
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (206) 622-4900
 ; TELEFAX: (206) 682-6031
 ; INFORMATION FOR SEQ ID NO: 13:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 318 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: linear
 ; US-09-098-079-13

Query Match 88.5%; Score 131; DB 9; Length 318;
 Best Local Similarity 90.3%; Pred. No. 1.2e-10;
 Matches 28; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ANLLLMVPIILAMAFMLTKILGYIQPR 31
 DB 4 ANLLLLVPIILAMAFMLTKILGYNQLR 34

RESULT 3
 US-10-428-487-34
 ; Sequence 34, Application US/10428487
 ; Publication No. US20040006780A1
 ; GENERAL INFORMATION:
 ; APPLICANT: RASTELLI, LUCA K.
 ; APPLICANT: GERBER, HANS-PETER
 ; TITLE OF INVENTION: VEGF-MODULATED GENES AND METHODS EMPLOYING THEM
 ; FILE REFERENCE: 0980080-0103
 ; CURRENT APPLICATION NUMBER: US/10/428,487
 ; CURRENT FILING DATE: 2003-05-02
 ; PRIOR APPLICATION NUMBER: 09/815,153
 ; PRIOR FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: 60/191,201
 ; PRIOR FILING DATE: 2000-03-22
 ; NUMBER OF SEQ ID NOS: 84
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 34
 ; LENGTH: 318
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-428-487-34

Query Match 88.5%; Score 131; DB 15; Length 318;
 Best Local Similarity 90.3%; Pred. No. 1.2e-10;

Matches 28; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ANLLLMVPIILAMAFMLTKILGYIQPR 31
 DB 4 ANLLLLVPIILAMAFMLTKILGYNQLR 34
 RESULT 4
 US-10-408-765A-189
 ; Sequence 189, Application US/10408765A
 ; Publication No. US20040101874A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ghosh, Soumitra S.
 ; APPLICANT: Fahy, Eoin D.
 ; APPLICANT: Zhang, Bing
 ; APPLICANT: Gibson, Bradford W.
 ; APPLICANT: Taylor, Steven W.
 ; APPLICANT: Glenn, Gary M.
 ; APPLICANT: Warnock, Dale E.
 ; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
 ; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
 ; FILE REFERENCE: 660088.465
 ; CURRENT APPLICATION NUMBER: US/10/408,765A
 ; CURRENT FILING DATE: 2003-04-04
 ; NUMBER OF SEQ ID NOS: 3077
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 189
 ; LENGTH: 318
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-408-765A-189

Query Match 88.5%; Score 131; DB 16; Length 318;
 Best Local Similarity 90.3%; Pred. No. 1.2e-10;
 Matches 28; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ANLLLMVPIILAMAFMLTKILGYIQPR 31
 DB 4 ANLLLLVPIILAMAFMLTKILGYNQLR 34

RESULT 5
 US-10-264-049-3171
 ; Sequence 3171, Application US/10264049
 ; Publication No. US20040005579A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Birse et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PA133PI
 ; CURRENT APPLICATION NUMBER: US/10/264,049
 ; CURRENT FILING DATE: 2002-10-04
 ; PRIOR APPLICATION NUMBER: PCT/US01/18569
 ; PRIOR FILING DATE: 2001-06-07
 ; PRIOR APPLICATION NUMBER: US 60/209,467
 ; PRIOR FILING DATE: 2000-06-07
 ; NUMBER OF SEQ ID NOS: 4360
 ; SOFTWARE: PatentIn Ver. 3.1
 ; SEQ ID NO 3171
 ; LENGTH: 101
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: MISC_FEATURE
 ; LOCATION: (1)
 ; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
 ; FEATURE:
 ; NAME/KEY: MISC_FEATURE
 ; LOCATION: (3)
 ; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
 ; FEATURE:
 ; NAME/KEY: MISC_FEATURE
 ; LOCATION: (4)
 ; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids

```
; FEATURE: MISC FEATURE
; NAME/KEY: MISC FEATURE
; LOCATION: (5)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
;
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (7)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
US-10-264-049-3171

Query Match      87.8%; Score 130; DB 15; Length 101;
Best Local Similarity 93.3%; Pred. No. 4.8e-11;
Matches 28; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy  2 NLLLVPIIIAMAFMLTKILGYIQPR 31
Db  21 NLLLVPIIIAMAFMLTKILGYIQLR 50

RESULT 6
US-10-425-115-310657
; Sequence 310657, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 310657
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Zea mays
;
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(261)
; OTHER INFORMATION: unsure at all Xaa locations
;
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_46377C.1.pep
US-10-425-115-310657

Query Match      87.8%; Score 130; DB 17; Length 261;
Best Local Similarity 93.3%; Pred. No. 1.3e-10;
Matches 28; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy  2 NLLLVPIIIAMAFMLTKILGYIQPR 31
Db  1 NLLLVPIIIAMAFMLTKILGYIQLR 30

RESULT 7
US-10-408-765A-1490
; Sequence 1490, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Bojin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1490
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-1490

Query Match      85.8%; Score 127; DB 16; Length 318;
Best Local Similarity 90.0%; Pred. No. 4.5e-10;
Matches 27; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy  2 NLLLVPIIIAMAFMLTKILGYIQPR 31
Db  5 NLLLVPIIIAMAFMLTKILGYIQLR 34

RESULT 8
US-10-264-049-3416
; Sequence 3416, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133PI
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 3416
; LENGTH: 79
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-264-049-3416

Query Match      81.1%; Score 120; DB 15; Length 79;
Best Local Similarity 92.9%; Pred. No. 1e-09;
Matches 26; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy  4 LLLVPIIIAMAFMLTKILGYIQPR 31
Db  1 LLLVPIIIAMAFMLTKILGYIQLR 28

RESULT 9
US-10-424-599-226150
; Sequence 226150, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 226150
; LENGTH: 110
; TYPE: PRT
; ORGANISM: Glycine max
;
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(110)
; OTHER INFORMATION: unsure at all Xaa locations
;
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_46243C.1.pep
US-10-424-599-226150
```

Query Match 75.3%; Score 111.5; DB 15; Length 110;
Best Local Similarity 84.4%; Pred. No. 2.5e-08;
Matches 27; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 1 ANLLLMVPIILIAAFML-TERKILGYIQPR 31
DB 17 APFLLLVPIILIAAFMLXTERKILGYIQLR 48

RESULT 10

US-10-282-122A-47875
; Sequence 47875, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent version 3.1
; SEQ ID NO 47875
; LENGTH: 355
; TYPE: PRT
; ORGANISM: Burkholderia cepacia
US-10-282-122A-47875

Query Match 43.9%; Score 65; DB 15; Length 355;
Best Local Similarity 41.4%; Pred. No. 0.4%;
Matches 12; Conservative 11; Mismatches 6; Indels 0; Gaps 0;

QY 3 LLLMVPIILIAAFML-TERKILGYIQPR 31
DB 28 ILVSVVILLCVAILLWERKILGMHVR 56

RESULT 11

US-10-425-115-332406
; Sequence 332406, Application US/10425115
; Publication No. US20040214272A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 332406
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(46)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_6626C.1.pep
US-10-425-115-332406

Query Match 43.2%; Score 64; DB 17; Length 46;
Best Local Similarity 82.4%; Pred. No. 0.069;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 13 AMAFLMLTERKILGYIQ 29

DB 1 AMAFLMLTERKILXYQ 17

RESULT 12

US-10-000-845-20
; Sequence 20, Application US/10000845
; Publication No. US20040197854A1
; GENERAL INFORMATION:
; APPLICANT: Brody, Howard
; Yaver, Deborah S.
; Lamsa, Michael
; Hansen, Kim
; TITLE OF INVENTION: Methods for Modifying the Production of
; a Polypeptide
; NUMBER OF SEQUENCES: 80
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Novo Nordisk of North America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/000,845
; FILING DATE: 24-Oct-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/928,692
; FILING DATE: 12-SEPT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4944.200-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 371 amino acids


```
;
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: None
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-10-000-845-20

Query Match      42.6%; Score 63; DB 17; Length 371;
Best Local Similarity 41.4%; Pred.No. 0.93;
Matches 12; Conservative 10; Mismatches 7; Indels 0; Gaps 0;

Qy  3 LLLLVPIILIAMAFMLTERKILGYIQPR 31
    :||:||||:||||:||||:||||:
Db  13 VLLVPSLLAVAVTVAERTKWTASMQRR 41

RESULT 13
US-10-425-115-202752
; Sequence 202752, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 202752
; LENGTH: 134
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(134)
; OTHER INFORMATION: unsure at all Xaa locations
; OTHER INFORMATION: Clone ID: MRT4577_11649C.1.pep
US-10-425-115-202752

Query Match      40.5%; Score 60; DB 17; Length 134;
Best Local Similarity 37.9%; Pred.No. 0.84;
Matches 11; Conservative 9; Mismatches 9; Indels 0; Gaps 0;

Qy  3 LLLLVPIILIAMAFMLTERKILGYIQPR 31
    :||:||||:||||:||||:||||:
Db  13 VLVTPVLLTVAFVTIAERTKWTASMQRR 41

RESULT 14
US-10-424-599-161174
; Sequence 161174, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 161174
; LENGTH: 138
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
```

```
;
; LOCATION: (1)..(138)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_116558C.1.pep
US-10-424-599-161174

Query Match      40.5%; Score 60; DB 15; Length 138;
Best Local Similarity 37.9%; Pred.No. 0.86;
Matches 11; Conservative 9; Mismatches 9; Indels 0; Gaps 0;

Qy  3 LLLLVPIILIAMAFMLTERKILGYIQPR 31
    :||:||||:||||:||||:||||:
Db  13 VLAVTPVLLTVAFVTIAERTKWTASMQRR 41

RESULT 15
US-10-282-122A-45946
; Sequence 45946, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA_034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 45946
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Bacillus anthracis
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (66)..(66)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (82)..(82)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC_FEATURE
```

```

; LOCATION: (197)..(197)
; OTHER INFORMATION: X=any amino acid
US-10-282-122A-45946

Query Match      37.2%; Score 55; DB 15; Length 261;
Best Local Similarity 54.5%; Pred. No. 9.1;
Matches 12; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY      5 LLMVPILLIAMAFLMLTERKILG 26
      ||| :||| :||| :|||
Db      112 LIMATTYVANGFLVLTNETRIIG 133

```

Search completed: November 11, 2004, 01:28:08
Job time : 34.3522 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 12:32:37 ; Search time 6.90846 Seconds
(without alignments)
191.991 Million cell updates/sec

Title: US-10-092-750-40

Perfect score: 106

Sequence: 1 SSQALRIHQWLHLFSDFTST 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2_6/ptodata/1/iaa/5A COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	47.2	156	4	US-09-248-796A-18650
2	46	43.4	1005	4	US-09-328-352-4877
3	45	42.5	56	4	US-09-621-976-5966
4	43.5	41.0	229	4	US-09-252-991A-18086
5	43	40.6	193	4	US-09-248-796A-18249
6	43	40.6	301	4	US-09-252-991A-24016
7	43	40.6	402	4	US-09-270-767-42272
8	43	40.6	534	4	US-09-293-549-8
9	42	39.6	594	4	US-09-107-532A-6929
10	41	38.7	63	4	US-09-248-796A-23739
11	41	38.7	93	4	US-09-248-796A-23391
12	41	38.7	115	4	US-09-513-999C-4944
13	41	38.7	216	3	US-08-464-052-6
14	41	38.7	216	3	US-08-461-002-6
15	41	38.7	216	3	US-08-689-411-6
16	41	38.7	216	4	US-08-392-210-6
17	41	38.7	240	4	US-09-270-767-33620
18	41	38.7	472	4	US-09-134-000C-5541
19	41	38.7	511	3	US-08-464-052-2
20	41	38.7	511	3	US-08-461-002-2
21	41	38.7	511	3	US-08-689-411-2
22	41	38.7	511	4	US-08-392-210-2
23	41	38.7	511	5	PCT-US94-09863-2
24	41	38.7	999	4	US-09-747-371-2
25	40.5	38.2	330	1	US-08-468-853-6
26	40.5	38.2	330	1	US-08-468-855-6
27	40.5	38.2	330	1	US-08-310-357-6

28	40.5	38.2	330	1	US-08-468-852-6
29	40.5	38.2	330	2	US-08-468-857-6
30	40	37.7	65	4	US-09-248-796A-27358
31	40	37.7	153	4	US-09-270-767-33889
32	40	37.7	153	4	US-09-270-767-49106
33	40	37.7	214	1	US-08-414-926A-23
34	40	37.7	214	1	US-08-926-922-23
35	40	37.7	214	3	US-09-253-682-23
36	40	37.7	214	3	US-09-527-657-23
37	40	37.7	214	4	US-09-892-100-23
38	40	37.7	222	4	US-09-252-991A-16816
39	40	37.7	585	4	US-09-489-039A-10279
40	39.5	37.3	87	4	US-09-248-796A-27519
41	39.5	37.3	770	4	US-09-107-532A-4388
42	39	36.8	108	4	US-09-270-767-35656
43	39	36.8	108	4	US-09-270-767-50873
44	39	36.8	224	4	US-09-270-767-45599
45	39	36.8	276	4	US-09-503-456-37

ALIGNMENTS

RESULT 1

US-09-248-796A-18650

; Sequence 18650, Application US/09248796A

; Patent No. 6747137

; GENERAL INFORMATION:

; APPLICANT: Keith Weinstock et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS

; FILE REFERENCE: 107196.132

; CURRENT APPLICATION NUMBER: US/09/248,796A

; CURRENT FILING DATE: 1999-02-12

; PRIOR APPLICATION NUMBER: US 60/074,725

; PRIOR FILING DATE: 1998-02-13

; PRIOR APPLICATION NUMBER: US 60/096,409

; PRIOR FILING DATE: 1998-08-13

; NUMBER OF SEQ ID NOS: 28208

; SEQ ID NO 18650

; LENGTH: 156

; TYPE: PRT

; ORGANISM: Candida albicans

US-09-248-796A-18650

Query Match 47.2%; Score 50; DB 4; Length 156;
Best Local Similarity 52.6%; Pred. No. 1.3;
Matches 10; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 SSQALRIHQWLHLFSDFTS 19

Db 71 SPQVTHSLHIFKNFTS 89

RESULT 2

US-09-328-352-4877

; Sequence 4877, Application US/09328352

; Patent No. 6562958

; GENERAL INFORMATION:

; APPLICANT: Gary L. Breton et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER

; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS

; CURRENT APPLICATION NUMBER: US/09/328,352

; CURRENT FILING DATE: 1999-06-04

; NUMBER OF SEQ ID NOS: 8252

; SEQ ID NO 4877

; LENGTH: 1005

; TYPE: PRT

; ORGANISM: Acinetobacter baumannii

US-09-328-352-4877

Query Match 43.4%; Score 46; DB 4; Length 1005;

```

; Best Local Similarity 53.8%; Pred. No. 42;
; Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 SQALRIHOWLHLF 14
Db 691 SQGFHVHQLAMF 703

RESULT 3
US-09-621-976-5966
; Sequence 5966, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621.976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 5966
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-621-976-5966

Query Match 42.5%; Score 45; DB 4; Length 56;
Best Local Similarity 56.7%; Pred. No. 2.7;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 SQALRIHOWLHL 13
Db 4 SQATRTSHWLXL 15

RESULT 4
US-09-252-991A-18086
; Sequence 18086, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18086
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18086

Query Match 41.0%; Score 43.5; DB 4; Length 229;
Best Local Similarity 41.7%; Pred. No. 22;
Matches 10; Conservative 5; Mismatches 4; Indels 5; Gaps 1;

QY 2 SQALRIHOWL-----HLFSDFTST 20
Db 149 AQALRYQLWLLALAFFAAWST 172

RESULT 5
US-09-248-796A-18249
; Sequence 18249, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:

```

```

; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248.796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 18249
; LENGTH: 193
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-18249

Query Match 40.6%; Score 43; DB 4; Length 193;
Best Local Similarity 55.6%; Pred. No. 22;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 8 HOWLHLFSD 16
Db 71 HEWVHLYDD 79

RESULT 6
US-09-252-991A-24016
; Sequence 24016, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24016
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24016

Query Match 40.6%; Score 43; DB 4; Length 301;
Best Local Similarity 77.8%; Pred. No. 35;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 QALRIHOWL 11
Db 76 QLLRLHOWL 84

RESULT 7
US-09-270-767-42272
; Sequence 42272, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7336-094
; CURRENT APPLICATION NUMBER: US/09/270.767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 42272
; LENGTH: 402
; TYPE: PRT
; ORGANISM: Drosophila melanogaster

```

US-09-270-767-42272

Query Match 40.6%; Score 43; DB 4; Length 402;
Best Local Similarity 56.2%; Pred. No. 48;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 4 ALRIHQWLHLSDFST 19
Db 271 SLLHLKSLHLPFTS 286

RESULT 8

US-09-293-549-8
Sequence 8, Application US/09293549

Patent No. 6440409

GENERAL INFORMATION:

APPLICANT: G. Todd Milne

APPLICANT: Gerald Fink

TITLE OF INVENTION: A METHOD FOR EFFICIENT AND HIGHLY

SELECTIVE CONTROL OF MICROORGANISMS

FILE REFERENCE: 50078/008002

CURRENT APPLICATION NUMBER: US/09/293,549

CURRENT FILING DATE: 1998-04-16

EARLIER APPLICATION NUMBER: 60/082,089

EARLIER FILING DATE: 1998-04-17

NUMBER OF SEQ ID NOS: 14

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 8

LENGTH: 534

TYPE: PRT

ORGANISM: Saccharomyces cerevisiae

US-09-293-549-8

Query Match 40.6%; Score 43; DB 4; Length 534;
Best Local Similarity 50.0%; Pred. No. 65;
Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Qy 1 SSQALRIHQWLHLSDFST 20
Db 482 SASALQNRKWLGLGDFDST 501

RESULT 9

US-09-107-532A-6929
Sequence 6929, Application US/09107532A

Patent No. 6583275

GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO

ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS

NUMBER OF SEQUENCES: 7310

CORRESPONDENCE ADDRESS:

ADDRESSEE: GENOME THERAPEUTICS CORPORATION

STREET: 100 Beaver Street

CITY: Waltham

STATE: Massachusetts

COUNTRY: USA

ZIP: 02354

COMPUTER READABLE FORM:

MEDIUM TYPE: CD-ROM ISO9660

COMPUTER: PC

OPERATING SYSTEM: <Unknown>

SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/107,532A

FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/085,598

FILING DATE: 14 May 1998

APPLICATION NUMBER: 60/051571

FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Ariniello, Pamela Denise

REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012

TELECOMMUNICATION INFORMATION:

TELEPHONE: (781)893-5007

TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 6929:

SEQUENCE CHARACTERISTICS:

LENGTH: 594 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: YES

ORIGINAL SOURCE:

ORGANISM: Enterococcus faecium

FEATURE:

NAME/KEY: misc feature

LOCATION: (B) LOCATION 1...594

SEQUENCE DESCRIPTION: SEQ ID NO: 6929:

US-09-107-532A-6929

Query Match 39.6%; Score 42; DB 4; Length 594;
Best Local Similarity 46.7%; Pred. No. 1.1e+02;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 6 RIHQWLHLSDFST 20
Db 9 KIRKYPHEFSNPLPT 23

RESULT 10

US-09-248-796A-23739

Sequence 23739, Application US/09248796A

Patent No. 6747137

GENERAL INFORMATION:

APPLICANT: Keith Weinstock et al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS

FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 107196.132

CURRENT APPLICATION NUMBER: US/09/248,796A

CURRENT FILING DATE: 1999-02-12

PRIOR APPLICATION NUMBER: US 60/074,725

PRIOR FILING DATE: 1998-02-13

PRIOR APPLICATION NUMBER: US 60/096,409

PRIOR FILING DATE: 1998-08-13

NUMBER OF SEQ ID NOS: 28208

SEQ ID NO 23739

LENGTH: 63

TYPE: PRT

ORGANISM: Candida albicans

US-09-248-796A-23739

Query Match 38.7%; Score 41; DB 4; Length 63;
Best Local Similarity 58.3%; Pred. No. 14;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 SSQALRIHQWLH 12
Db 15 SSAILNLHLWLH 26

RESULT 11

US-09-248-796A-23391

Sequence 23391, Application US/09248796A

Patent No. 6747137

GENERAL INFORMATION:

APPLICANT: Keith Weinstock et al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS

FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 107196.132

CURRENT APPLICATION NUMBER: US/09/248,796A

CURRENT FILING DATE: 1999-02-12

PRIOR APPLICATION NUMBER: US 60/074,725

PRIOR FILING DATE: 1998-02-13

```
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 23391
; LENGTH: 93
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-23391

Query Match          38.7%; Score 41; DB 4; Length 93;
Best Local Similarity 53.3%; Pred. No. 21;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 5 LRIHOWLHLSFSDFTS 19
   :||:|:|:|:|:|
Db 73 LNVHQEFLVFSNFTS 87

RESULT 12
US-09-513-999C-4944
; Sequence 4944, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 4944
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 25
; OTHER INFORMATION: Xaa=Pro or Thr
US-09-513-999C-4944

Query Match          38.7%; Score 41; DB 4; Length 115;
Best Local Similarity 31.6%; Pred. No. 26;
Matches 6; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 SSQALRIHOWLHLSFSDFTS 19
   :||:|:|:|:|:|
Db 94 ASRIWNLQFTQLYKDI TS 112

RESULT 13
US-08-464-052-6
; Sequence 6, Application US/08464052
; Patent No. 6008201
; GENERAL INFORMATION:
; APPLICANT: Riley M.D., Lee W.
; TITLE OF INVENTION: DNA Molecule Encoding for Cellular
; TITLE OF INVENTION: Uptake of Mycobacterium Tuberculosis and Uses thereof
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle
; STREET: Clinton Square, P.O. Box 1051
; CITY: Rochester
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,002
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldman, Michael L.
; REGISTRATION NUMBER: 30,727
; REFERENCE/DOCKET NUMBER: 19603/186 (D-1485B)
; TELEPHONE: (716) 263-1304
; TELEFAX: (716) 263-1600
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 216 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-461-002-6

Query Match          38.7%; Score 41; DB 3; Length 216;
Best Local Similarity 47.1%; Pred. No. 51;
Matches 8; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 4 ALRIHOWLHLSFSDFTST 20
   :||:|:|:|:|
Db 29 AQRVMDLHPDGLDIT 45

RESULT 14
US-08-461-002-6
; Sequence 6, Application US/08461002
; Patent No. 6214543
; GENERAL INFORMATION:
; APPLICANT: Riley M.D., Lee W.
; TITLE OF INVENTION: DNA Molecule Encoding for Cellular
; TITLE OF INVENTION: Uptake of Mycobacterium Tuberculosis and Uses thereof
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle
; STREET: Clinton Square, P.O. Box 1051
; CITY: Rochester
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,002
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldman, Michael L.
; REGISTRATION NUMBER: 30,727
; REFERENCE/DOCKET NUMBER: 19603/186 (D-1485B)
; TELEPHONE: (716) 263-1304
; TELEFAX: (716) 263-1600
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 216 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-461-002-6

Query Match          38.7%; Score 41; DB 3; Length 216;
Best Local Similarity 47.1%; Pred. No. 51;
Matches 8; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 4 ALRIHOWLHLSFSDFTST 20
   :||:|:|:|:|
Db 29 AQRVMDLHPDGLDIT 45
```

OY 4 ALRIHQWLHFSDFST 20
 | : | | | |
 DB 29 AQRVMDWLHPDGLTDT 45

RESULT 15
 US-08-689-411-6
 ; Sequence 6, Application US/08689411
 ; Patent No. 6224881
 ; GENERAL INFORMATION:
 ; APPLICANT: Riley M.D., Lee W.
 ; APPLICANT: Chong, Pele
 ; TITLE OF INVENTION: DNA MOLECULE FRAGMENTS ENCODING FOR
 ; TITLE OF INVENTION: CELLULAR UPTAKE OF MYCOBACTERIUM TUBERCULOSIS AND USES
 ; TITLE OF INVENTION: THEREOF
 ; NUMBER OF SEQUENCES: 14
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
 ; STREET: Clinton Square, P.O. Box 1051
 ; CITY: Rochester
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 14603
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/689,411
 ; FILING DATE:
 ; CLASSIFICATION: 536
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Goldman, Michael L.
 ; REGISTRATION NUMBER: 30,727
 ; REFERENCE/DOCKET NUMBER: 19603/187
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (716) 263-1304
 ; TELEFAX: (716) 263-1600
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 216 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: peptide
 ; US-08-689-411-6

Query Match 38.7%; Score 41; DB 3; Length 216;
 Best Local Similarity 47.1%; Pred. No. 51;
 Matches 8; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

OY 4 ALRIHQWLHFSDFST 20
 | : | | | |
 DB 29 AQRVMDWLHPDGLTDT 45

Search completed: November 10, 2004, 13:44:10
 Job time : 6.90846 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 16:36:12 ; Search time 21.4853 Seconds
(without alignments)

328.807 Million cell updates/sec

Title: US-10-092-750-40

Perfect score: 106

Sequence: 1 SSQALRIHQWLHLFSDFTST 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	106	100.0	20	US-10-092-750-40	Sequence 40, Appl
2	47	44.3	132	US-10-425-115-317501	Sequence 317501,
3	47	44.3	675	US-10-302-267-42	Sequence 42, Appl
4	47	44.3	675	US-10-412-699B-704	Sequence 704, Appl
5	46	43.4	369	US-10-104-047-2357	Sequence 2357, Ap
6	45.5	42.9	175	US-10-425-115-341865	Sequence 341865
7	45	42.5	238	US-09-727-855B-9	Sequence 9, Appl
8	45	42.5	335	US-10-425-114-38355	Sequence 38355, A
9	45	42.5	460	US-10-425-114-70131	Sequence 70131, A
10	44	41.5	83	US-10-425-115-359581	Sequence 359581,
11	44	41.5	126	US-10-425-115-345444	Sequence 345444,
12	44	41.5	173	US-10-424-599-232253	Sequence 232253,
13	44	41.5	184	US-10-424-599-235750	Sequence 235750,

14	44	41.5	284	16	US-10-408-765A-2745	Sequence 2745, Ap
15	44	41.5	475	15	US-10-424-599-249784	Sequence 249784,
16	44	41.5	503	14	US-10-093-352-5	Sequence 5, Appl
17	44	41.5	503	14	US-10-093-352-43	Sequence 43, Appl
18	43.5	41.0	106	16	US-10-437-963-171215	Sequence 171215,
19	43	40.6	52	15	US-10-424-599-271059	Sequence 271059,
20	43	40.6	60	14	US-10-011-585A-185	Sequence 185, App
21	43	40.6	70	17	US-10-425-115-302525	Sequence 302525,
22	43	40.6	76	9	US-09-764-877-1576	Sequence 1576, Ap
23	43	40.6	76	15	US-10-242-515-1576	Sequence 1576, Ap
24	43	40.6	92	16	US-10-437-963-169776	Sequence 169776,
25	43	40.6	166	15	US-10-424-599-229308	Sequence 229308,
26	43	40.6	182	16	US-10-767-701-51233	Sequence 51233, A
27	43	40.6	290	17	US-10-739-930-5913	Sequence 5913, Ap
28	43	40.6	474	17	US-10-425-115-240066	Sequence 240066,
29	43	40.6	498	14	US-10-369-493-12692	Sequence 12692, A
30	43	40.6	501	14	US-10-099-352-44	Sequence 44, Appl
31	43	40.6	562	17	US-10-425-115-240067	Sequence 240067,
32	43	40.6	832	16	US-10-437-963-147028	Sequence 147028,
33	43	40.6	981	17	US-10-425-115-349629	Sequence 349629,
34	42.5	40.1	103	17	US-10-425-115-344352	Sequence 344352,
35	42	39.6	46	17	US-10-425-115-226935	Sequence 226935,
36	42	39.6	60	16	US-10-437-963-183557	Sequence 183557,
37	42	39.6	70	16	US-10-437-963-112327	Sequence 112327,
38	42	39.6	77	17	US-10-425-115-230987	Sequence 230987,
39	42	39.6	94	15	US-10-284-049-3360	Sequence 3360, Ap
40	42	39.6	119	17	US-10-425-115-264604	Sequence 264604,
41	42	39.6	124	15	US-10-424-599-176499	Sequence 176499,
42	42	39.6	417	16	US-10-437-963-145688	Sequence 145688,
43	42	39.6	473	14	US-10-178-782-5	Sequence 5, Appl
44	42	39.6	473	14	US-10-178-782-6	Sequence 6, Appl
45	42	39.6	586	15	US-10-282-122A-77600	Sequence 77600, A

ALIGNMENTS

RESULT 1
US-10-092-750-40
; Sequence 40, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; APPLICANT: Wright, Martin C.
; TITLE OF INVENTION: Polypeptides Interactive with BCL-XI
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092,750
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/274,526
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40
; TYPE: PRT
; LENGTH: 20
; ORGANISM: Homo sapiens
US-10-092-750-40

Query Match 100.0%; Score 106; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.4e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSQALRIHQWLHLFSDFTST 20
Db 1 SSQALRIHQWLHLFSDFTST 20

RESULT 2
US-10-425-115-317501
; Sequence 317501, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:

```
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovacic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425.115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 317501
; LENGTH: 132
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MET4577_52639C.1.pap
US-10-425-115-317501

Query Match 44.3%; Score 47; DB 17; Length 132;
Best Local Similarity 87.5%; Pred. No. 17;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 8 HQWLHLS 15
DB 102 HQWLHFT 109

RESULT 3
US-10-302-267-42
; Sequence 42, Application US/10302267
; Publication No. US20030229915A1
; GENERAL INFORMATION:
; APPLICANT: Keddle, James
; APPLICANT: Fromm, Michael
; APPLICANT: Heard, Jacqueline
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Adam, Luc
; APPLICANT: Broun, Pierre
; APPLICANT: Pineda, Omaira
; APPLICANT: Reuber, Lynne
; APPLICANT: Zhang, James
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Samaha, Raymond
; APPLICANT: Creelman, Robert
; APPLICANT: Creelman, Robert
; TITLE OF INVENTION: PLANT GENE SEQUENCES II
; FILE REFERENCE: MBI-0007
; CURRENT APPLICATION NUMBER: US/10/302,267
; CURRENT FILING DATE: 2002-11-22
; PRIOR FILING DATE: US/09/506,720
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/120,880
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: 60/121,037
; PRIOR FILING DATE: 1999-02-22
; PRIOR APPLICATION NUMBER: 60/124,278
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: 60/129,450
; PRIOR FILING DATE: 1999-04-15
; PRIOR APPLICATION NUMBER: 60/135,134
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: 60/144,153
; PRIOR FILING DATE: 1999-07-15
; PRIOR APPLICATION NUMBER: 60/161,143
; PRIOR FILING DATE: 1999-10-22
; PRIOR APPLICATION NUMBER: 60/162,656
; PRIOR FILING DATE: 1999-11-01
; NUMBER OF SEQ ID NOS: 218
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 42
; LENGTH: 675
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G1055
US-10-412-699B-704
```

```
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G1055
US-10-302-267-42

Query Match 44.3%; Score 47; DB 14; Length 675;
Best Local Similarity 58.3%; Pred. No. 85;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 SSQALRIHQWLH 12
DB 510 SSEGRRLHQWFH 521

RESULT 4
US-10-412-699B-704
; Sequence 704, Application US/10412699B
; Publication No. US20040045049A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: Zhang, James
; APPLICANT: Fromm, Michael E.
; APPLICANT: Heard, Jacqueline E.
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Adam, Luc J.
; APPLICANT: Broun, Pierre E.
; APPLICANT: Pineda, Omaira
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Keddle, James S.
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Samaha, Raymond R.
; APPLICANT: Filgrim, Marsha L.
; APPLICANT: Creelman, Robert A.
; APPLICANT: DuBell, Arnold N.
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Kumimoto, Roderick
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: Polynucleotides and Polypeptides in Plants
; FILE REFERENCE: MBI-0048CIP
; CURRENT APPLICATION NUMBER: US/10/412,699B
; CURRENT FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: 09/394,519
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: 09/489,376
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: 09/506,720
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 09/533,030
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,392
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,029
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/532,591
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,648
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/713,994
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: 09/819,142
; PRIOR FILING DATE: 2001-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2011
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 704
; LENGTH: 675
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G1055
US-10-412-699B-704
```

```
Query Match      44.3%; Score 47; DB 15; Length 675;
Best Local Similarity 58.3%; Pred. No. 85;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 SSQALRIHOWLH 12
   |||:|:|:|:|
Db 510 SSEGKRLHQWPH 521

RESULT 5
US-10-104-047-2357
; Sequence 2357, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20030236392A1el full length cDNA
; FILE REFERENCE: HI-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2357
; LENGTH: 369
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-2357

Query Match      43.4%; Score 46; DB 14; Length 369;
Best Local Similarity 61.5%; Pred. No. 67;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 QALRIHOWLHFS 15
   |||:|:|:|
Db 17 RALRSHRWLSFS 29

RESULT 6
US-10-425-115-341865
; Sequence 341865, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 341865
; LENGTH: 175
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_74948C.1.pep
US-10-425-115-341865

Query Match      42.9%; Score 45.5; DB 17; Length 175;
Best Local Similarity 55.0%; Pred. No. 38;
Matches 11; Conservative 3; Mismatches 5; Indels 1; Gaps 1;

QY 1 SSQA-LRIHOWLHFSDFTS 19
   |||:|:|:|:|
Db 151 SNQAVRWGGWLLLFCDPCS 170

RESULT 7
US-09-727-855B-9
; Sequence 9, Application US/09727855B
```

```
; Patent No. US20020168703A1
; GENERAL INFORMATION:
; APPLICANT: HOSHINO, Tatsuo
; APPLICANT: OJIMA, Kazuyuki
; APPLICANT: SETOGUCHI, Yufaka
; TITLE OF INVENTION: PROCESS FOR THE MANUFACTURE OF CAROTENOIDS AND BIOLOGICALLY USEFUL
; FILE REFERENCE: C38435/111694
; CURRENT APPLICATION NUMBER: US/09/727,855B
; CURRENT FILING DATE: 2000-12-01
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Phaffia rhodozyma
US-09-727-855B-9

Query Match      42.5%; Score 45; DB 9; Length 238;
Best Local Similarity 80.0%; Pred. No. 52;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 7 IHQWLHFLFS 16
   |||:|:|:|
Db 74 VHQVLEHFLFS 83

RESULT 8
US-10-425-114-38355
; Sequence 38355, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 38355
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700051738_FLI.pep
US-10-425-114-38355

Query Match      42.5%; Score 45; DB 15; Length 335;
Best Local Similarity 57.1%; Pred. No. 87;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 QALRIHOWLHFLFS 16
   |||:|:|:|
Db 113 QALDLHQHVHVPFD 126

RESULT 9
US-10-425-114-70131
; Sequence 70131, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
```

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(5313)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 70131
 ; LENGTH: 460
 ; TYPE: PRT
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: UC-ZMFLB73042B05_FLI.pep
 US-10-425-114-70131

Query Match 42.5%; Score 45; DB 15; Length 460;
 Best Local Similarity 57.1%; Pred. No. 1.2e+02;
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 QALRIHOWLHLPD 16
 DB 104 QALDLHQHVHVPD 117

RESULT 10
 US-10-425-115-359581
 ; Sequence 359581, Application US/10425115
 ; Publication No. US20040214272A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(53222)B
 ; CURRENT APPLICATION NUMBER: US/10/425,115
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 369326
 ; SEQ ID NO 359581
 ; LENGTH: 83
 ; TYPE: PRT
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (1)..(83)
 ; OTHER INFORMATION: unsure at all Xaa locations
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: MRT4577_91109C.1.pep
 US-10-425-115-359581

Query Match 41.5%; Score 44; DB 17; Length 83;
 Best Local Similarity 58.3%; Pred. No. 31;
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 6 RIHOWLHLPD 17
 DB 53 RSXDLHVHVPD 64

RESULT 11
 US-10-425-115-345444
 ; Sequence 345444, Application US/10425115
 ; Publication No. US20040214272A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(53222)B
 ; CURRENT APPLICATION NUMBER: US/10/425,115
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 369326

; SEQ ID NO 345444
 ; LENGTH: 126
 ; TYPE: PRT
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: MRT4577_78212C.1.pep
 US-10-425-115-345444

Query Match 41.5%; Score 44; DB 17; Length 126;
 Best Local Similarity 47.1%; Pred. No. 48;
 Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 SOALRIHOWLHLPD 18
 DB 32 SVALENNRWHFCSPT 48

RESULT 12
 US-10-424-599-232253
 ; Sequence 232253, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 232253
 ; LENGTH: 173
 ; TYPE: PRT
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT3847_51748C.1.pep
 US-10-424-599-232253

Query Match 41.5%; Score 44; DB 15; Length 173;
 Best Local Similarity 50.0%; Pred. No. 65;
 Matches 9; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 3 QALRIHOWLHLPD 20
 DB 111 QCLYHSLWHDGFFVT 128

RESULT 13
 US-10-424-599-235750
 ; Sequence 235750, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 235750
 ; LENGTH: 184
 ; TYPE: PRT
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (1)..(184)
 ; OTHER INFORMATION: unsure at all Xaa locations
 ; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT3847_54910C.1.pep
US-10-424-599-235750

Query Match 41.5%; Score 44; DB 15; Length 184;
Best Local Similarity 63.6%; Pred. No. 69;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 SSQALRIHOWLH 12
|:|:|:|:|:|:
DB 70 SSQALRISEWMH 80

RESULT 14

US-10-408-765A-2745
; Sequence 2745, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2745
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-2745

Query Match 41.5%; Score 44; DB 16; Length 284;
Best Local Similarity 66.7%; Pred. No. 1.le+02;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 SSQALRIHOWLH 12
|:|:|:|:|:|:
DB 155 STGALRIHGWTH 166

RESULT 15

US-10-424-599-249784
; Sequence 249784, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 249784
; LENGTH: 475
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(475)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_67584C.1.pep
US-10-424-599-249784

Query Match 41.5%; Score 44; DB 15; Length 475;
Best Local Similarity 47.1%; Pred. No. 1.8e+02;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 SSQALRIHOWLHLPDSF 17
|:|:|:|:|:|:|:|:|:|:|:|:
DB 189 SSEAFRIHEAIFCAGP 205

Search completed: November 11, 2004, 01:28:22
Job time : 22.5353 secs

This Page Blank (uspio)

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	46	43.4	272	2	S52977		hypothetical prote
2	46	43.4	275	2	D84769		hypothetical prote
3	45	42.5	264	2	T40103		hypothetical prote
4	44	41.5	185	2	A70075		hypothetical prote
5	44	41.5	1276	2	T03204		probable tail-host
6	44	41.5	1291	2	T03273		probable tail-host
7	43.5	41.0	164	2	D83172		hypothetical prote
8	43	40.6	139	2	E82419		conserved hypothe
9	43	40.6	332	2	A90436		coA-ligase / coenz
10	43	40.6	424	2	C82594		hypothetical prote
11	43	40.6	443	2	D82975		two-component sens
12	43	40.6	468	2	AD2395		two-component sens
13	43	40.6	534	2	S34679		diphthamide synth
14	42	39.6	218	2	H69005		hypothetical prote
15	42	39.6	253	2	A73159		protein-1-isoaspar
16	42	39.6	292	2	G97325		transcription regu
17	42	39.6	309	2	S52256		copper export prot
18	42	39.6	586	2	D83484		Sgar protein VCA02
19	42	39.6	750	2	AG3008		polyketide synthet
20	42	39.6	770	2	F98275		satramycin mx1 syn
21	41	38.7	144	2	S69309		hypothetical prote
22	41	38.7	230	2	JC7972		spermatogenesis-re
23	41	38.7	317	2	F70566		hypothetical prote
24	41	38.7	324	2	F63504		proliferating-cell
25	41	38.7	340	2	B88339		protein C05E4.11 [
26	41	38.7	370	2	T09918		hypothetical prote
27	41	38.7	396	2	D64022		hypothetical prote
28	41	38.7	423	2	AH0702		conserved hypothe
29	41	38.7	465	2	S73854		hypothetical prote

Query Match 40.6%; Score 43; DB 2; Length 339;
Best Local Similarity 55.6%; Pred. No. 26;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 8 HQWLHLFSD 16
|:|:|:|:|:
Db 171 HEWLHIWTD 179

RESULT 10
C82594
hypothetical protein XP2151 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: C82594
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: C82594
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-424 <SIM>
A:Cross-references: UNIPROT:Q9PBJ3; GB:AE004029; GB:AE003849; NID:g9107276; PIDN:AAF848954
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; AJ
B:Rios, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H.
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
Submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Prohme
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigret
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E.
A:Authors: Martins, E.M.F.; Mateukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasaka
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Za
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XP2151

Query Match 40.6%; Score 43; DB 2; Length 424;
Best Local Similarity 53.8%; Pred. No. 34;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 8 HQWLHLFSDTST 20
|:|:|:|:|:
Db 164 HPYLHLFVSST 176

RESULT 11
D82975
two-component sensor Phor PA5361 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: D82975
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Foiger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
Nature 406, 959-964, 2000
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: D82975
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-443 <STO>
A:Cross-references: UNIPROT:P23621; GB:AE004948; GB:AE004091; NID:g9951680; PIDN:AAG0874
A:Experimental source: strain PA01
C:Genetics:
A:Gene: phoR; PA5361
C:Superfamily: phosphate regulon regulatory protein; sensor histidine kinase homology

Query Match 40.6%; Score 43; DB 2; Length 443;
 Best Local Similarity 77.8%; Pred. No. 35;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 QALRIHOWL 11
 |||
 Db 52 QLLRLHOWL 60
 |||

RESULT 12
 AD2395
 C:Species: Nostoc sp. PCC 7120
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
 C:Accession: AD2395
 R:Kaneke, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Shimizu, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
 DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120
 A:Reference number: AB1807; MUID:21595285; PMID:11759840
 A:Accession: AD2395
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-468 <KUR>
 A:Cross-references: UNIPROT:Q8YN54; GB:BA000019; PIDN:BA076415.1; PID:g17133853; GSPDB:G
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: alr4716

Query Match 40.6%; Score 43; DB 2; Length 468;
 Best Local Similarity 87.5%; Pred. No. 37;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 10 WLHLFSDF 17
 |||
 Db 29 WLHLSDP 36
 |||

RESULT 13
 S34679
 diphthamide synthesis protein DPH2 - yeast (Saccharomyces cerevisiae)
 N:Alternate names: protein YKL191w
 C:Species: Saccharomyces cerevisiae
 A:Variety: strain S288C
 C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
 C:Accession: B59228; S33961; JN0855; S38023; S38028; S52281; S30773; S34679; S49568
 R:Wiemann, S.; Voss, H.; Schwager, C.; Rupp, T.; Stegemann, J.; Zimmermann, J.; Grothues
 J. Bacteriol. 179, 1715-1755, 1997
 A:Title: Sequencing and analysis of 51.6 kilobases on the left arm of chromosome XI from
 A:Reference number: S49568; MUID:94205264; PMID:9154185
 A:Accession: B59228
 A>Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-534 <WIE>
 A:Cross-references: UNIPROT:P32461; EMBL:X74151; NID:9450365; PIDN:CAA52247.1; PID:g3952
 A:Note: the published sequence was revised in GenBank
 R:Cheret, G.; Mattheakis, L.C.; Sor, F.
 Yeast 9, 661-667, 1993
 A:Title: DNA sequence analysis of the YCN2 region of chromosome XI in Saccharomyces cere
 A:Reference number: S33960; MUID:93348776; PMID:8394042
 A:Accession: S33961
 A:Molecule type: DNA
 A:Residues: 1-534 <CHE>
 A:Cross-references: GB:X69765; NID:g296985; PIDN:CAA49420.1; PID:g296987
 R:Mattheakis, L.C.; Sor, F.; Collier, R.O.
 Gene 132, 149-154, 1993
 A:Title: Diphthamide synthesis in Saccharomyces cerevisiae: Structure of the DPH2 gene.
 A:Reference number: JN0855; MUID:94010339; PMID:8406038
 A:Accession: JN0855
 A:Molecule type: DNA
 A:Residues: 1-534 <MAT>

A:Cross-references: GB:L01424; NID:g171413; PIDN:AAA64990.1; PID:g765109
 R:Wiemann, S.; Voss, H.; Schwager, C.; Rupp, T.; Grothues, D.; Sensen, C.; Stegemann, J.;
 submitted to the Protein Sequence Database, March 1994
 A:Reference number: S37825
 A:Accession: S38023
 A:Molecule type: DNA
 A:Residues: 1-534 <WIE>
 A:Cross-references: EMBL:Z28191; NID:g486338; PIDN:CAA82035.1; PID:g486339; MIPS:YKL191w
 R:Maia e Silva, A.; Bossier, P.; Villela, C.; Fernandes, L.; Soares, H.; Guerreiro, P.; R
 submitted to the Protein Sequence Database, March 1994
 A:Reference number: S38024
 A:Accession: S38028
 A:Molecule type: DNA
 A:Residues: 1-534 <MAI>
 A:Cross-references: EMBL:Z28191; NID:g486338; PIDN:CAA82035.1; PID:g486339; MIPS:YKL191w
 R:Pardo, J.M.
 submitted to the EMBL Data Library, September 1993
 A:Description: The protein phosphatase calcineurin is essential for NaCl tolerance in Sac
 A:Reference number: S52281
 A:Accession: S52281
 A:Molecule type: DNA
 A:Residues: 258-534 <PAR>
 A:Cross-references: EMBL:Z26521; NID:g473144; PIDN:CAA81289.1; PID:g473145
 C:Genetics:
 A:Gene: SGD:DPH2
 A:Cross-references: SGD:S0001674; MIPS:YKL191w
 A:Map position: 11L
 C:Keywords: diphthamide biosynthesis

Query Match 40.6%; Score 43; DB 2; Length 534;
 Best Local Similarity 50.0%; Pred. No. 43;
 Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 SSQALRIHQWLHFLSDFTST 20
 |||
 Db 482 SASALQNRWKGSGSDPDST 501
 |||

RESULT 14
 H69005
 hypothetical protein MTH1040 - Methanobacterium thermoautotrophicum (strain Delta H)
 C:Species: Methanobacterium thermoautotrophicum
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
 C:Accession: H69005
 R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; F
 ; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwan, N.;
 ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
 J. Bacteriol. 179, 7135-7155, 1997
 A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: functi
 A:Reference number: A69000; MUID:98037514; PMID:9371463
 A:Accession: H69005
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-218 <MTH>
 A:Cross-references: UNIPROT:O27119; GB:AE000876; GB:AE000666; NID:g2622140; PIDN:AAB8553
 A:Experimental source: strain Delta H
 C:Genetics:
 A:Gene: MTH1040
 A:Start codon: GTG

Query Match 39.6%; Score 42; DB 2; Length 218;
 Best Local Similarity 55.6%; Pred. No. 24;
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 8 HOWLHFLSD 16
 |||
 Db 174 HWLHVISE 182
 |||

RESULT 15
 A75159
 protein-L-isoaspartate methyltransferase homolog PAB0283 - Pyrococcus abyssi (strain Orsa
 C:Species: Pyrococcus abyssi

C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
R;anonymous, Genoscope
A;Accession: A75159
A;Submitted to the EMBL Data Library, July 1999
A;Description: *Pyrococcus abyssi* Genome sequence: insights into archaeal chromosome structure
A;Reference number: A75001
A;Accession: A75159
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-253 <YAW>
A;Cross-references: UNIPROT:Q9V1J7; GB:AJ248284; GB:AL096836; NID:g5457730; PIDN:CAB4935
A;Experimental source: strain Orsay
C;Genetics:
A;Gene: pimt-like; PAB0283

Query Match 39.6%; Score 42; DB 2; Length 253;
Best Local Similarity 41.2%; Pred. No. 28;
Matches 7; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

```
QY      1 SSQALRIHQWLHLSDF 17
         | : | : | : | : | :
Db     197 SNQVRLHEKLEFKDY 213
```

Search completed: November 10, 2004, 13:40:48
Job time : 5.45596 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 12:27:34 ; Search time 24.4214 Seconds
(without alignments)
471.205 Million cell updates/sec

Title: US-10-092-750-40
Perfect score: 106
Sequence: 1 SSQALRIHQWLHLFSDFTST 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters:	1825181
--	---------

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_02:*

```
1: uniprot_sprot:*
2: uniprot_trembl:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query %			DB	ID	Description
	Score	Match	Length			
1	51	48.1	680	2	Q9PSZ0	Q9PSZ0 pseudopleur
2	51	48.1	974	2	Q93Z68	Q93Z68 pseudopleur h
3	49	46.2	1078	2	Q9DG19	Q9DG19 meleagrid h
4	48	45.3	192	2	Q8SVU0	Q8SVU0 encephallito
5	48	45.3	420	2	Q8BI96	Q8BI96 shewanella
6	47	44.3	78	2	Q84SK4	Q84SK4 oryza sativ
7	47	44.3	205	2	Q7PP27	Q7PP27 arapheles g
8	47	44.3	326	2	Q8W029	Q8W029 arabidopsis
9	47	44.3	481	2	Q8PGM7	Q8PGM7 xanthomonas
10	47	44.3	675	2	Q8GHU4	Q8GHU4 arabidopsis
11	47	44.3	675	2	Q9SG86	Q9SG86 arabidopsis
12	46	43.4	272	1	YCR3 ERWHE	Q01334 erwinia her
13	46	43.4	353	2	Q9U3D3	Q9U3D3 caenorhabdi
14	46	43.4	387	2	Q82290	Q82290 arabidopsis
15	46	43.4	1581	2	Q6SUZ0	Q6SUZ0 debaryomyce
16	46	43.4	2127	2	Q8URG9	Q8URG9 australian
17	46	43.4	2369	2	Q67940	Q67940 canis fami
18	46	43.4	2369	2	Q67941	Q67941 canis fami
19	46	43.4	2369	2	Q67942	Q67942 canis fami
20	46	43.4	2369	2	AA887834	AA887834 canis fam
21	46	43.4	2369	2	AA887835	AA887835 canis fam
22	46	43.4	2369	2	AA887836	AA887836 canis fam
23	45	42.5	184	2	Q6ZUP8	Q6ZUP8 homo sapien
24	45	42.5	184	2	BAC86171	BAC86171 homo sapi
25	45	42.5	264	2	Q74796	Q74796 schizosacch
26	45	42.5	474	2	Q8Y308	Q8Y308 halstonia s
27	44	41.5	185	1	YXEG BACSU	YXEG BACSU bacillus su
28	44	41.5	191	2	Q73QK9	Q73QK9 treponema d
29	44	41.5	191	2	AA510929	AA510929 treponema
30	44	41.5	229	2	Q7U3M1	Q7U3M1 synechococc
31	44	41.5	463	2	Q86HB9	Q86HB9 dictyosteli

32	44	41.5	503	2	Q9UTU1	Q9UTU1
33	44	41.5	504	2	Q9C148	Q9C148
34	44	41.5	870	2	Q7RH85	Q7RH85
35	44	41.5	886	2	Q8IJ31	Q8IJ31
36	44	41.5	910	2	Q37959	Q37959
37	44	41.5	1070	2	Q8EGF3	Q8EGF3
38	44	41.5	1276	2	Q6A284	Q6A284
39	44	41.5	1291	2	Q64295	Q64295
40	44	41.5	2671	2	Q75433	Q75433
41	44	41.5	2671	2	AAS53540	AAS53540
42	43.5	41.0	164	2	Q9HXU5	Q9HXU5
43	43.5	41.0	757	1	AOC2_MOUSE	AOC2_MOUSE
44	43	40.6	92	1	VAPD_NEIGO	VAPD_NEIGO
45	43	40.6	92	2	Q6ZK08	Q6ZK08
46	43	40.6	92	2	Q6ZK08	Q6ZK08
47	43	40.6	92	2	Q6ZK08	Q6ZK08
48	43	40.6	92	2	Q6ZK08	Q6ZK08
49	43	40.6	92	2	Q6ZK08	Q6ZK08
50	43	40.6	92	2	Q6ZK08	Q6ZK08
51	43	40.6	92	2	Q6ZK08	Q6ZK08
52	43	40.6	92	2	Q6ZK08	Q6ZK08
53	43	40.6	92	2	Q6ZK08	Q6ZK08
54	43	40.6	92	2	Q6ZK08	Q6ZK08
55	43	40.6	92	2	Q6ZK08	Q6ZK08
56	43	40.6	92	2	Q6ZK08	Q6ZK08
57	43	40.6	92	2	Q6ZK08	Q6ZK08
58	43	40.6	92	2	Q6ZK08	Q6ZK08
59	43	40.6	92	2	Q6ZK08	Q6ZK08
60	43	40.6	92	2	Q6ZK08	Q6ZK08
61	43	40.6	92	2	Q6ZK08	Q6ZK08
62	43	40.6	92	2	Q6ZK08	Q6ZK08
63	43	40.6	92	2	Q6ZK08	Q6ZK08
64	43	40.6	92	2	Q6ZK08	Q6ZK08
65	43	40.6	92	2	Q6ZK08	Q6ZK08
66	43	40.6	92	2	Q6ZK08	Q6ZK08
67	43	40.6	92	2	Q6ZK08	Q6ZK08
68	43	40.6	92	2	Q6ZK08	Q6ZK08
69	43	40.6	92	2	Q6ZK08	Q6ZK08
70	43	40.6	92	2	Q6ZK08	Q6ZK08
71	43	40.6	92	2	Q6ZK08	Q6ZK08
72	43	40.6	92	2	Q6ZK08	Q6ZK08
73	43	40.6	92	2	Q6ZK08	Q6ZK08
74	43	40.6	92	2	Q6ZK08	Q6ZK08
75	43	40.6	92	2	Q6ZK08	Q6ZK08
76	43	40.6	92	2	Q6ZK08	Q6ZK08
77	43	40.6	92	2	Q6ZK08	Q6ZK08
78	43	40.6	92	2	Q6ZK08	Q6ZK08
79	43	40.6	92	2	Q6ZK08	Q6ZK08
80	43	40.6	92	2	Q6ZK08	Q6ZK08
81	43	40.6	92	2	Q6ZK08	Q6ZK08
82	43	40.6	92	2	Q6ZK08	Q6ZK08
83	43	40.6	92	2	Q6ZK08	Q6ZK08
84	43	40.6	92	2	Q6ZK08	Q6ZK08
85	43	40.6	92	2	Q6ZK08	Q6ZK08
86	43	40.6	92	2	Q6ZK08	Q6ZK08
87	43	40.6	92	2	Q6ZK08	Q6ZK08
88	43	40.6	92	2	Q6ZK08	Q6ZK08
89	43	40.6	92	2	Q6ZK08	Q6ZK08
90	43	40.6	92	2	Q6ZK08	Q6ZK08
91	43	40.6	92	2	Q6ZK08	Q6ZK08
92	43	40.6	92	2	Q6ZK08	Q6ZK08
93	43	40.6	92	2	Q6ZK08	Q6ZK08
94	43	40.6	92	2	Q6ZK08	Q6ZK08
95	43	40.6	92	2	Q6ZK08	Q6ZK08
96	43	40.6	92	2	Q6ZK08	Q6Z

ALIGNMENTS

RESULT 1					
ID	Q9PSZO	PRELIMINARY;	PRT;	680 AA.	
AC	Q9PSZO;				
DT	01-MAY-2000 (T-EMBLrel. 13, Created)				
DT	01-MAY-2000 (T-EMBLrel. 13, Last sequence update)				
DT	01-MAR-2004 (T-EMBLrel. 26, Last annotation update)				
DE	Aminopeptidase N (EC 3.4.11.2) (Fragment).				
GN	Name=ampN;				
OS	Pseudopleuronectes americanus (Winter flounder) (Pleuronectes americanus).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;				
OC	Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;				
OC	Pleuronectoidae; Pleuronectidae; Pseudopleuronectes.				
OX	NBI_TaxID=8265;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Douglas S.E., Gallant J.W., Bullerwell C.E.;				
RT	"Molecular Investigation of Aminopeptidase N Expression in the Winter Flounder, Pleuronectes americanus.";				
RL	J. Appl. Ichthyol. 0:0-0(1998).				
DR	EMBL; AF043383; AAC32807.1; -.				
DR	MEROPS; M01.001; -.				
DR	GO; GO:0004177; F:aminopeptidase activity; IEA.				
DR	GO; GO:0016787; F:hydrolase activity; IEA.				
DR	GO; GO:0004179; F:membrane alanyl aminopeptidase activity; IEA.				
DR	GO; GO:0006508; P:proteolysis and peptidolysis; IEA.				
DR	InterPro; IPR019303; Peptidase_M1.				
DR	InterPro; IPR006025; Pept_Mzn_BS.				
DR	Fam; PF01433; Peptidase_M1.				
DR	PRINTS; PS00756; ALADIPTASE.				
DR	PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.				
KW	Aminopeptidase; Hydrolase.				
FT	NON TER 680				
SQ	SEQUENCE 680 AA; 75883 MW; 0D75D2ECB8E8FC8 CRC64;				
Query Match 48.1%; Score 51; DB 2; Length 680;					
Best Local Similarity 80.0%; Pred. No. 19;					
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps					
QY	9 QWLHFSDET 18				
Db	182 QWHLFTDFT 191				
RESULT 2					
ID	Q93268	PRELIMINARY;	PRT;	974 AA.	
AC	Q93268;				
DT	01-NOV-1998 (T-EMBLrel. 08, Created)				
DT	01-NOV-1998 (T-EMBLrel. 08, Last sequence update)				
DT	01-MAR-2004 (T-EMBLrel. 26, Last annotation update)				
DE	Aminopeptidase N (EC 3.4.11.2).				

GN Name=ampN;
 OS Pseudopleuronectes americanus (Winter flounder) (Pleuronectes americanus)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Nacteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes; Pleuronectoidae; Pleuronectidae; Pseudopleuronectes.
 OX NCBI_TaxID=8265;
 [1]
 RN SEQUENCE FROM N.A.
 RP TISSUE=Intestine;
 RC Douglas S.E., Gallant J.W., Bullerwell C.E.;
 RA "Molecular Investigation of Aminopeptidase N Expression in the Winter Flounder, *Pleuronectes americanus*."
 RL J. Appl. Ichthyol. 0:0-0(1998).
 DR EMBL; AF012465; AAC32754.1; -.
 DR MEROPS; M01.001; -.
 DR GO; GO:0004177; F:aminopeptidase activity; IEA.
 DR GO; GO:0016787; F:hydrolase activity; IEA.
 DR GO; GO:0004179; F:membrane alanyl aminopeptidase activity; IEA.
 DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR001930; Peptidase_M1.
 DR InterPro; IPR006025; Pept_M_Zn_BS.
 DR Pfam; PF01433; Peptidase_M1; 1.
 DR PRINTS; PR00756; ALADIPTASE.
 DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
 KW Aminopeptidase; Hydrolase.
 SQ SEQUENCE 974 AA; 109743 MW; BCE551C948892977 CRC64;
 Query Match 48.1%; Score 51; DB 2; Length 974;
 Best Local Similarity 80.0%; Pred. No. 27;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 9 QWHLFSDFT 18
 Db 182 QWHLFTDFT 191
 RESULT 3
 QSDGY9 PRELIMINARY; PRT; 1078 AA.
 AC QSDGY9
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE UL52 DNA helicase-primase associated protein (UL52 DNA helicase/primase complex protein).
 DE Name=HVT060; Synonyms=UL52;
 GN Meleagrid herpesvirus 1 (herpesvirus of turkeys).
 OS Viruses; dsDNA viruses, no RNA stage; Herpesviridae; Alphaherpesvirinae; Marek's disease-like viruses.
 OX NCBI_TaxID=37108;
 [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=FC126;
 RC MEDLINE=20578232; PubMed=11134310;
 RA Afonso C.L., Tulman E.R., Lu Z., Zsak L., Rock D.L., Kutish G.F.;
 RT "The genome of turkey herpesvirus."
 RL J. Virol. 75:971-978(2001).
 [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=FC126;
 RC MEDLINE=21195611; PubMed=11297687;
 RA Kingham B.F., Zelnik V., Kopacek J., Majerciak V., Ney E., Schmidt C.J.;
 RT "The genome of herpesvirus of turkeys: comparative analysis with Marek's disease viruses."
 RL J. Gen. Virol. 82:1123-1135(2001).
 DR EMBL; AF281866; AAG45790.1; -.
 DR EMBL; AF282130; AAG30093.1; -.
 DR GO; GO:0003896; F:DNA primase activity; IEA.
 DR GO; GO:0004386; F:helicase activity; IEA.
 DR GO; GO:0006260; P:DNA replication; IEA.

DR InterPro; IPR004340; UL52 UL70.
 DR Pfam; PF03121; Herpes_UL52; 1.
 KW Helicase.
 SQ SEQUENCE 1078 AA; 120951 MW; AA1F0CF511BDF65 CRC64;
 Query Match 46.2%; Score 49; DB 2; Length 1078;
 Best Local Similarity 47.1%; Pred. No. 64;
 Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
 QY 2 SQALRIHOWLHFLSDFT 18
 Db 715 SQSVFAEWLHFLTSEYT 731
 RESULT 4
 QBSVUO PRELIMINARY; PRT; 192 AA.
 AC QBSVUO
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hypothetical protein ECU04_0880.
 DE Name=ECU04_0880;
 GN Encephalitozoon cuniculi.
 OS Eukaryota; Fungi; Microsporidia; Unikaryonidae; Encephalitozoon.
 OC Eukaryota; Fungi; Microsporidia; Unikaryonidae; Encephalitozoon.
 OX NCBI_TaxID=6035;
 [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=GB-M1;
 RC MEDLINE=21576510; PubMed=11719806;
 RA Katinka M.D., Duprat S., Cornillot E., Metenier G., Thomarat P., Prensier G., Barbe V., Peyretailade E., Brottier P., Wincker P., Delbac F., El Alaoui H., Peyret P., Saurin W., Gouy M., Weissenbach J., Vivares C.P.;
 RA "Genome sequence and gene compaction of the eukaryote parasite Encephalitozoon cuniculi."
 RT Encephalitozoon cuniculi.
 RT Nature 414:450-453(2001).
 [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=GB-M1;
 RC Genoscope;
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL590444; CAD25275.1; -.
 DR GO; GO:0004852; F:uroporphyrinogen-III synthase activity; IEA.
 DR GO; GO:0006783; F:heme biosynthesis; IEA.
 DR InterPro; IPR003754; HEM4_synth.
 KW Hypothetical protein.
 SQ SEQUENCE 192 AA; 22306 MW; FAF2E41DD91D0B31 CRC64;
 Query Match 45.3%; Score 48; DB 2; Length 192;
 Best Local Similarity 46.7%; Pred. No. 15;
 Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
 QY 2 SQALRIHOWLHFLSD 16
 Db 43 SQGIRREWLHMFTE 57
 RESULT 5
 QSEI96 PRELIMINARY; PRT; 420 AA.
 AC QSEI96;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE ATP-dependent RNA helicase SrmB.
 DE Name=srmB; OrderedLocusNames=SO0947;
 GN Shewanella oneidensis.
 OS Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales; Alteromonadaceae; Shewanella.
 OX NCBI_TaxID=70863;
 [1]
 RN SEQUENCE FROM N.A.

```
RC STRAIN=MR-1;
RX MEDLINE=22297686; PubMed=1268813; DOI=10.1038/nbt749;
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA Read T.D., Eisen J.A., Sehadi R., Ward N.L., Meche B.A.,
RA Clayton R.A., Meyer T., Teapin A., Scott J., Bearan M.J.,
RA Brinkac L.M., Daugherty S.C., DeBoy R.T., Dodson R.J., Durkin A.S.,
RA Hatt D.H., Kolonay J.F., Madupu R., Peterson J.D., Umayam L.A.,
RA White O., Wolf A.M., Vamathevan J.J., Weidman J.F., Imbraim M.,
RA Lee K., Berry K.J., Lee C., Mueller J., Khouri H.M., Gill J.,
RA Utterback T.R., McDonald L.A., Feldblyum T.V., Smith H.O.,
RA Venter J.C., Neilson K.H., Fraser C.M.;
RA "Genome sequence of the dissimilatory metal ion-reducing bacterium
RT Shewanella oneidensis.";
RL Nat. Biotechnol. 20:1118-1123(2002).
CC -!- SIMILARITY: Belongs to the DEAD box helicase family.
DR EMBL; AE015540; AAN54021.1; -.
DR HSSP; P10081; 1QDE.
DR TIGR; SO0947; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008036; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F:hydrolyase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR0011410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000629; DEAD_box_N.
DR InterPro; IPR001650; Helicase_C.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELICG; 1.
DR PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
DR ATP-binding; Complete Proteome; Helicase; Hydrolase.
KW ATP-binding; Complete Proteome; Helicase; Hydrolase.
SQ SEQUENCE 420 AA; 46809 MW; C5F83898314E38F8 CRC64;

Query Match 45.3%; Score 48; DB 2; Length 420;
Best Local Similarity 63.6%; Pred. No. 34;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 6 RIHOWLHFPD 16
Db :|||||
218 KIHQWIHLADD 228

RESULT 6
Q84SK4 PRELIMINARY; PRT; 78 AA.
AC Q84SK4
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE P0014E04.8 protein.
GN Name=P0014E04.8
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
CX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Saeki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
RA Wu J., Nilmura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,
RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,
RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,
RA Hijishita S., Honda M., Ichikawa Y., Igonuma A., Iijima M., Ikeda M.,
RA Ikeno M., Itoh S., Itoh Y., Itoh Y., Itoh Y., Iwabuchi A., Kamiya K.,
RA Karasawa M., Katagiri S., Kiuta A., Kobayashi N., Kono I.,
RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,
RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,
RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai, K., Shibata M.,
RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,
RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,
RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,
RA Yano M., Jiang J., Gojobori T.;
```

```
RT "The genome sequence and structure of rice chromosome 1.";
RL Nature 420:312-316(2002).
DR EMBL; AF004362; BAC55975.1; -.
DR Gramene; Q84SK4; -.
SQ SEQUENCE 78 AA; 8527 MW; 4112738F1B770850 CRC64;

Query Match 44.3%; Score 47; DB 2; Length 78;
Best Local Similarity 50.0%; Pred. No. 8;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 4 ALRIHOWLHLPDSFTS 19
Db :|||||
40 SLLHSHWLCFLRFAS 55

RESULT 7
Q7PP27 PRELIMINARY; PRT; 205 AA.
AC Q7PP27
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ENSANGP00000021244 (Fragment).
GN Name=ENSANGP00000018755;
OS Anopheles gambiae str. FST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
CX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAAB01008960; EAA11128.2; -.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; ehand; 3.
DR ProDom; PD000012; EF-hand; 2.
DR PROSITE; PS00018; EF_HAND; UNKNOWN_3.
FT NON TER
FT 1
SQ SEQUENCE 205 AA; 23197 MW; 12DE4512C1ECB614 CRC64;

Query Match 44.3%; Score 47; DB 2; Length 205;
Best Local Similarity 63.6%; Pred. No. 23;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 7 IHOWLHLPDSF 17
Db :|||||
105 INQWLHVFPNP 115

RESULT 8
Q8W029 PRELIMINARY; PRT; 326 AA.
AC Q8W029
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE AtbZIP transcription factor (fragment).
GN Name=AtbZIP28;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
CX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21481677; PubMed=11597504;
RA Stracks R., Weber M., Weisshaar B.;
RA "The R2R3-MYB gene family in Arabidopsis thaliana.";
```

```

RL Curr. Opin. Plant Biol. 4:447-456(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20243261; PubMed=10785665;
RA Stracke R., Werber M., Weisshaar B.;
RT "The WRKY superfamily of plant transcription factors.";
RL Trends Plant Sci. 5:199-206(2000).
RN [3]
RP SEQUENCE FROM N.A.
RA Jakoby M.J.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ419850; CAD12033.1; -
FT NON TER 1
SQ SEQUENCE 326 AA; 34980 MW; 6A60B3C155BE9D9C CRC64;

Query Match 44.3%; Score 47; DB 2; Length 326;
Best Local Similarity 58.3%; Pred. No. 37;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 SSQALRIHQWLH 12
|||:|||||
Db 161 SSEGRLHQWFH 172

RESULT 9
Q8PGM7 PRELIMINARY; PRT; 481 AA.
AC Q8PGM7;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 26, Last annotation update)
DE Integral membrane protein.
OS OrderedLocustNames=XAC3589;
GN Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonas
OC Xanthomonadaceae; Xanthomonas.
CX NCBI_TaxID=92829;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=2202145; PubMed=12024217;
RA Da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Canarotte G., Cannavan F., Cardoso J., Chantreiro F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.B., Cursino-Santos J.R., El-Dorry H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White P.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities";
RL Nature 417:459-463(2002).
DR EMBL; AE012009; AAM38432.1; -
GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR000150; Hypothet.coif.
DR InterPro; IPR000537; UbiA_prenyltrans.
DR Pfam; PF01040; UbiA; 1.
DR PROSITE; PS01228; COF_1; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 481 AA; 52904 MW; C31C59563D29069A CRC64;

Query Match 44.3%; Score 47; DB 2; Length 481;
Best Local Similarity 37.9%; Pred. No. 57;
Matches 11; Conservative 2; Mismatches 4; Indels 12; Gaps 1;

QY 3 QALRIHQWL-----HLFSDFTS 19
|||:|||||
Db 198 KALRIHQWLKLVFVPLLTAAHRLDLES 226

RESULT 10
Q8GUH4 PRELIMINARY; PRT; 675 AA.
AC Q8GUH4;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative bZIP transcription factor.
GN Name=At3g10800;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
CX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Southwick A., Nguyen M., Tripp M., Palm C.J., Jones T., Wu T.,
RA Carninci P., Chen H., Cheuk R., Chan M.M., Chang C.H., Dale J.M.,
RA Deng J.M., Hayashizaki Y., Hsuan V.W., Lee J.M., Ishida J., Kamiya A.,
RA Kawai J., Kim C.J., Narusaka M., Onodera C.S., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Shinn P., Tang C.C., Toroumi M., Wong C., Wu H.C.,
RA Yamada K., Yu G., Yuan S., Shinozaki K., Ecker J., Theologis A.,
RA Davis R.W.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
RC -!- SIMILARITY: Belongs to the bZIP family.
DR EMBL; BT002502; AAC00862.1; -
GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR008917; Euk.transcr_DNA.
DR InterPro; IPR004827; TF_bZIP.
DR Pfam; PF00170; bZIP; 1.
DR SMART; SM00338; BRLZ; 1.
DR PROSITE; PS0217; bZIP; 1.
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 675 AA; 73460 MW; 953B29B12A762F85 CRC64;

Query Match 44.3%; Score 47; DB 2; Length 675;
Best Local Similarity 58.3%; Pred. No. 82;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 SSQALRIHQWLH 12
|||:|||||
Db 510 SSEGRLHQWFH 521

RESULT 11
Q9SG86 PRELIMINARY; PRT; 675 AA.
AC Q9SG86;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative bZIP transcription factor.
GN Name=T7M13.12;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
CX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Lin X., Kaul S., Town C.D., Benito M.-I., Creasy T.H., Haas B.,
RA Rensing C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RC -!- SIMILARITY: Belongs to the bZIP family.
DR EMBL; AC011708; AAF19569.1; -
GO; GO:0005634; C:nucleus; IEA.

```



```

DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR008917; Euk transr_DNA.
DR InterPro; IPR004827; TF_bZip.
DR Pfam; PF00170; bZip; 1.
DR SMART; SM00338; BRL2; 1.
DR PROSITE; PS50217; bZip; 1.
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 675 AA; 73432 MW; DAA74401C581C078 CRC64;

Query Match 44.3%; Score 47; DB 2; Length 675;
Best Local Similarity 58.3%; Pred. No. 82;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 SSQALRIHQWLH 12
   ||: ||: ||: ||
Db 510 SSEGKRLHQWPF 521

RESULT 12
YCR3_ERWHE
ID_YCR3_ERWHE STANDARD; PRT; 272 AA.
AC Q01334;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Hypothetical 29.9 kDa protein in crtE 3'region (ORF3).
OS Erwinia herbicola.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pantoea.
OX NCBI_TaxID=549;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=EH010;
RX MEDLINE=95107236; PubMed=7808389;
RA Hundle B., Alberti M., Nivelestein V., Beyer P., Kleinig H.,
RA Armstrong G.A., Burke D.H., Hearst J.E.;
RT "Functional assignment of Erwinia herbicola Eh010 carotenoid genes
RT expressed in Escherichia coli.";
RL Mol. Gen. Genet. 245:406-416(1994).
CC -1- SIMILARITY: Belongs to the sodium:galactoside symporter family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M87280; AAA64975.1; -.
DR PIR; S52977; S52977.
DR InterPro; IPR001927; Na/Gal symport.
DR PROSITE; PS00872; NA_GALACTOSIDE_SYMP; FALSE_NEG.
KW Hypothetical protein.
SQ SEQUENCE 272 AA; 29955 MW; C7C753B416F14AB6 CRC64;

Query Match 43.4%; Score 46; DB 1; Length 272;
Best Local Similarity 75.0%; Pred. No. 45;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 QALRIHQWLHLP 14
   ||||| ||: ||: ||
Db 13 QALRQRLWLCLF 24

RESULT 13
Q9U3D3
ID_Q9U3D3 PRELIMINARY; PRT; 353 AA.
AC_Q9U3D3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

```

RA Ishida J., Jiang P.X., Jones T., Kaniya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC005314; AAC36171.2; -;
DR EMBL; AY074560; AAL67100.1; -;
DR EMBL; AF412061; AAL06514.1; -;
DR PIR; D84769; D84769.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004765; F:shikimate Kinase activity; IEA.
DR GO; GO:0008652; P:amino acid biosynthesis; IEA.
DR InterPro; IPR008978; HSP20 chap.
DR InterPro; IPR000623; Shik_Kinase.
DR Pfam; PF01202; SKI; 1.
DR SEQUENCE 387 AA; 42687 MW; CF784D442E76503A CRC64;
SQ
Query Match 43.4%; Score 46; DB 2; Length 387;
Best Local Similarity 61.5%; Pred. No. 65;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 6 RIHOWLHLFSDFT 18
DB 271 RADQWRHLYSGFT 283

RESULT 15
Q6BUZO
ID Q6BUZO PRELIMINARY; PRT; 1581 AA.
AC Q6BUZO;
DT 01-OCT-2004 (TrEMBLrel. 28, Created)
DT 01-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Similar to CA1156|CarLr1 Candida albicans.
GN ORFNames=DEHAOC07601g;
OS Debaryomyces hansenii (Yeast) (Torulaspora hansenii).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Debaryomycetes.
OX NCBI_TaxID=4959;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CBST767;
EG GENOLEVURES;
RA DuJon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Talla E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Baray S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boisrame A., Boyer J., Cattolico L., Confanioli F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantaye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Niclaud J.M., Nikolski M., Ostas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.F., Straub M.D., Suleau A.,
RA Swennene D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Winkler P., Souciet J.L.;
RT "Genome evolution in yeasts."
RL Nature 430:35-44 (2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CBST767;
EG GENOSCODE;
RA Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR382135; CAG86037.1; -;
DR SEQUENCE 1581 AA; 182145 MW; DEFF4AAB263B32E8 CRC64;
SQ
Query Match 43.4%; Score 46; DB 2; Length 1581;
Best Local Similarity 58.3%; Pred. No. 3e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 6 RIHOWLHLFSDFT 17
DB 536 RKHEWLYFRNF 547

Search completed: November 10, 2004, 13:38:46
JOB time : 26.4214 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 12:32:37 ; Search time 6.21762 Seconds
(without alignments)
191.991 Million cell updates/sec

Title: US-10-092-750-41

Perfect score: 90

Sequence: 1 QGVGRQLAIIGDDINRRK 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/iaa/5A COMB pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B COMB pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A COMB pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B COMB pep.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS COMB pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1 pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	85	94.4	19	3	US-09-236-385A-35
2	85	94.4	20	3	US-09-236-385A-36
3	85	94.4	27	4	US-09-544-664B-4
4	85	94.4	28	1	US-08-440-391-2
5	85	94.4	28	1	US-08-440-391-18
6	85	94.4	28	2	US-08-308-597A-2
7	85	94.4	28	2	US-08-308-597A-18
8	85	94.4	28	3	US-09-236-385A-2
9	85	94.4	28	3	US-09-236-385A-18
10	85	94.4	28	5	PCT-US96-06122-2
11	85	94.4	28	5	PCT-US96-06122-18
12	85	94.4	36	1	US-08-440-391-14
13	85	94.4	36	2	US-08-308-597A-14
14	85	94.4	36	3	US-09-236-385A-14
15	85	94.4	36	5	PCT-US96-06122-14
16	85	94.4	117	4	US-09-381-488-4
17	85	94.4	141	1	US-08-471-058-23
18	85	94.4	152	1	US-08-471-058-22
19	85	94.4	210	3	US-08-471-057-22
20	85	94.4	210	4	US-08-470-865-22
21	85	94.4	211	1	US-08-321-071A-16
22	85	94.4	211	1	US-08-471-058-7
23	85	94.4	211	1	US-08-471-058-9
24	85	94.4	211	1	US-08-471-058-10
25	85	94.4	211	1	US-08-471-058-11
26	85	94.4	211	2	US-08-944-530-2
27	85	94.4	211	2	US-08-944-530-4

28 85 94.4 211 3 US-08-471-057-7 Sequence 7, Appli
29 85 94.4 211 3 US-08-471-057-9 Sequence 9, Appli
30 85 94.4 211 3 US-08-471-057-10 Sequence 10, Appli
31 85 94.4 211 3 US-08-471-057-11 Sequence 11, Appli
32 85 94.4 211 4 US-09-381-488-2 Sequence 2, Appli
33 85 94.4 211 4 US-08-470-865-7 Sequence 7, Appli
34 85 94.4 211 4 US-08-470-865-9 Sequence 9, Appli
35 85 94.4 211 4 US-08-470-865-10 Sequence 10, Appli
36 85 94.4 211 4 US-08-470-865-11 Sequence 11, Appli
37 85 94.4 211 4 US-09-155-327G-13 Sequence 13, Appli
38 83 92.2 27 4 US-09-544-664B-5 Sequence 5, Appli
39 80 88.9 16 4 US-09-544-664B-30 Sequence 30, Appli
40 80 88.9 17 4 US-09-544-664B-57 Sequence 57, Appli
41 80 88.9 19 4 US-10-196-080-2 Sequence 2, Appli
42 79 87.8 31 1 US-08-440-391-3 Sequence 3, Appli
43 79 87.8 31 1 US-08-440-391-16 Sequence 16, Appli
44 79 87.8 31 2 US-08-908-597A-3 Sequence 3, Appli
45 79 87.8 31 2 US-08-908-597A-16 Sequence 16, Appli

ALIGNMENTS

RESULT 1
US-09-236-385A-35
; Sequence 35, Application US/09236385A
; Patent No. 6221615
; GENERAL INFORMATION:
; APPLICANT: CHITTENDEN, Thomas D.; and
; LUTZ, Robert J.
; TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH
; MODULATE APOPTOSIS
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hale and Dorr
; STREET: 1455 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/236,385A
; FILING DATE: 25-Jan-1999
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: WIXON, HENRY N.
; REGISTRATION NUMBER: 32,073
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-942-8400
; TELEFAX: 202-942-8484
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 35:
US-09-236-385A-35

Query Match 94.4%; Score 85; DB 3; Length 19;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGVGRQLAIIGDDINRR 17
DB 2 QGVGRQLAIIGDDINRR 18

```
RESULT 2
US-09-236-385A-36
; Sequence 36, Application US/09236385A
; Patent No. 6221615
; GENERAL INFORMATION:
; APPLICANT: CHITTENDEN, Thomas D.; and
; LUTZ, Robert J.
; TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH
; MODULATE APOPTOSIS
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hale and Dorr
; STREET: 1455 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/236.385A
; FILING DATE: 25-Jan-1999
; CLASSIFICATION: <unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: WIXON, HENRY N.
; REGISTRATION NUMBER: 32,073
; (C) ATTORNEY DOCKET NO. 104322.147CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-942-8400
; TELEFAX: 202-942-8484
; INFORMATION FOR SEQ ID NO: 36
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 36
US-09-236-385A-36
Query Match 94.4%; Score 85; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.6e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GQVGRQLAIIGDDINRR 17
Db 3 GQVGRQLAIIGDDINRR 19

RESULT 3
US-09-544-664B-4
; Sequence 4, Application US/09544664B
; Patent No. 6713280
; GENERAL INFORMATION:
; APPLICANT: Huang, Ziwei
; APPLICANT: Wang, Jialun
; APPLICANT: Zhang, Zhijia
; APPLICANT: Shan, Simel
; APPLICANT: Lu, Zhixian
; TITLE OF INVENTION: Enhancement of Peptide Cellular Uptake
; FILE REFERENCE: 8321-68
; CURRENT APPLICATION NUMBER: US/09/544.664B
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/US00/09352
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/128,202
; PRIOR FILING DATE: 1999-04-07
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 27
; TYPE: PRI
```

```
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Peptide
; OTHER INFORMATION: segment from BH3 domain of a Bcl-2 superfamily
; OTHER INFORMATION: polypeptide
US-09-544-664B-4
Query Match 94.4%; Score 85; DB 4; Length 27;
Best Local Similarity 100.0%; Pred. No. 2.2e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GQVGRQLAIIGDDINRR 17
Db 6 GQVGRQLAIIGDDINRR 22

RESULT 4
US-08-440-391-2
; Sequence 2, Application US/08440391
; Patent No. 5656725
; GENERAL INFORMATION:
; APPLICANT: CHITTENDEN, Thomas D.; and
; APPLICANT: LUTZ, Robert J.
; TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH
; MODULATE APOPTOSIS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hale and Dorr
; STREET: 1455 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/440.391
; FILING DATE: 12-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: WIXON, HENRY N.
; REGISTRATION NUMBER: 32,073
; REFERENCE/DOCKET NUMBER: 104322.147
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-942-8400
; TELEFAX: 202-942-8484
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 amino acid
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-440-391-2
Query Match 94.4%; Score 85; DB 1; Length 28;
Best Local Similarity 100.0%; Pred. No. 2.3e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GQVGRQLAIIGDDINRR 17
Db 6 GQVGRQLAIIGDDINRR 22

RESULT 5
US-08-440-391-18
; Sequence 18, Application US/08440391
; Patent No. 5656725
; GENERAL INFORMATION:
; APPLICANT: CHITTENDEN, Thomas D.; and
; APPLICANT: LUTZ, Robert J.
; TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH
```

```

; TITLE OF INVENTION: MODULATE APOPTOSIS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hale and Dorr
; STREET: 1455 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 12-MAY-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: WIXON, HENRY N.
; REGISTRATION NUMBER: 32,073
; REFERENCE/DOCKET NUMBER: 104322.147
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-942-8400
; TELEFAX: 202-942-8484
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: Peptide
; US-08-440-391-18

Query Match 94.4%; Score 85; DB 1; Length 28;
Best Local Similarity 100.0%; Pred. No. 2.3e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GQVGRQLAIIIGDDINRR 17
Db 6 GQVGRQLAIIIGDDINRR 22

RESULT 6
US-08-908-597A-2
; Sequence 2, Application US/08908597A
; Patent No. 5863795
; GENERAL INFORMATION:
; APPLICANT: CHITTENDEN, Thomas D.; and
; APPLICANT: LUTZ, Robert J.
; TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH
; TITLE OF INVENTION: MODULATE APOPTOSIS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hale and Dorr
; STREET: 1455 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION NUMBER: US/08/440.391
; FILING DATE: 12-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: WIXON, HENRY N.
; REGISTRATION NUMBER: 32,073
; REFERENCE/DOCKET NUMBER: 104322.147
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-942-8400
; TELEFAX: 202-942-8484
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-908-597A-18

Query Match 94.4%; Score 85; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 2.3e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GQVGRQLAIIIGDDINRR 17
Db 6 GQVGRQLAIIIGDDINRR 22

RESULT 7
US-08-908-597A-18
; Sequence 18, Application US/08908597A
; Patent No. 5863795
; GENERAL INFORMATION:
; APPLICANT: CHITTENDEN, Thomas D.; and
; APPLICANT: LUTZ, Robert J.
; TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH
; TITLE OF INVENTION: MODULATE APOPTOSIS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hale and Dorr
; STREET: 1455 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION NUMBER: US/08/440.391
; FILING DATE: 12-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: WIXON, HENRY N.
; REGISTRATION NUMBER: 32,073
; REFERENCE/DOCKET NUMBER: 104322.147
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-942-8400
; TELEFAX: 202-942-8484
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-908-597A-18

Query Match 94.4%; Score 85; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 2.3e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GQVGRQLAIIIGDDINRR 17
Db 6 GQVGRQLAIIIGDDINRR 22

RESULT 8
```

us-10-092-750-41.ra1

Fri Nov 12 14:55:24 2004

```

US-09-236-385A-2
; Sequence 2, Application US/09236385A
; Patent No. 6221615
; GENERAL INFORMATION:
; APPLICANT: CHITTENDEN, Thomas D.; and
; LUTZ, Robert J.
; TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH
; MODULATE APOPTOSIS
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hale and Dorr
; STREET: 1455 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/236,385A
; FILING DATE: 25-Jan-1999
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: WIXON, HENRY N.
; REGISTRATION NUMBER: 32,073
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-942-8400
; TELEFAX: 202-942-8484
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 amino acid
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-236-385A-2

```

```

Query Match          94.4%; Score 85; DB 3; Length 28;
Best Local Similarity 100.0%; Pred. No. 2.3e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 GQVGRQLAIIIGDDINRR 17
        |||||
Db      6 GQVGRQLAIIIGDDINRR 22

```

```

RESULT 9
US-09-236-385A-18
; Sequence 18, Application US/09236385A
; Patent No. 6221615
; GENERAL INFORMATION:
; APPLICANT: CHITTENDEN, Thomas D.; and
; LUTZ, Robert J.
; TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH
; MODULATE APOPTOSIS
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hale and Dorr
; STREET: 1455 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/236,385A
; FILING DATE: 25-Jan-1999

```

```

; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: WIXON, HENRY N.
; REGISTRATION NUMBER: 32,073
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-942-8400
; TELEFAX: 202-942-8484
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-09-236-385A-18

```

```

Query Match          94.4%; Score 85; DB 3; Length 28;
Best Local Similarity 100.0%; Pred. No. 2.3e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 GQVGRQLAIIIGDDINRR 17
        |||||
Db      6 GQVGRQLAIIIGDDINRR 22

```

```

RESULT 10
PCT-US96-06122-2
; Sequence 2, Application PC/TUS9606122
; GENERAL INFORMATION:
; APPLICANT: IMMUNOGEN, INC.
; TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS
; WHICH MODULATE APOPTOSIS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hale and Dorr
; STREET: 1455 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/06122
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRICE APPLICATION DATA:
; APPLICATION NUMBER: US 08/440,391
; FILING DATE: 12-MAY-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: WIXON, HENRY N.
; REGISTRATION NUMBER: 32,073
; REFERENCE/DOCKET NUMBER: 104322.147PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-942-8400
; TELEFAX: 202-942-8484
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 amino acid
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
PCT-US96-06122-2

```

```

Query Match          94.4%; Score 85; DB 5; Length 28;
Best Local Similarity 100.0%; Pred. No. 2.3e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 GQVGRQLAIIIGDDINRR 17
        |||||

```

```
Db      6 GQVGRQLAIGDDINRR 22
|||||
PCT-US96-06122-18
; Sequence 18, Application PC/TUS9606122
; GENERAL INFORMATION:
; APPLICANT: IMMUNOGEN, INC.
; TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS
; WHICH MODULATE APOPTOSIS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hale and Dorr
; STREET: 1455 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: PCT/US96/06122
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/440,391
; FILING DATE: 12-MAY-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: WIXON, HENRY N.
; REGISTRATION NUMBER: 32,073
; REFERENCE/DOCKET NUMBER: 104322.147PCT
; TELEPHONE: 202-942-8400
; TELEFAX: 202-942-8484
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US96-06122-18

Query Match      94.4%; Score 85; DB 5; Length 28;
Best Local Similarity 100.0%; Pred. No. 2.3e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GQVGRQLAIGDDINRR 17
|||||
Db      6 GQVGRQLAIGDDINRR 22
|||||

RESULT 12
US-08-440-391-14
; Sequence 14, Application US/08440391
; Patent No. 5656725
; GENERAL INFORMATION:
; APPLICANT: CHITTENDEN, Thomas D.; and
; APPLICANT: LUTZ, Robert J.
; TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH
; MODULATE APOPTOSIS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hale and Dorr
; STREET: 1455 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

US-08-908-597A-14
; Sequence 14, Application US/08908597A
; Patent No. 5863795
; GENERAL INFORMATION:
; APPLICANT: CHITTENDEN, Thomas D.; and
; APPLICANT: LUTZ, Robert J.
; TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH
; MODULATE APOPTOSIS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hale and Dorr
; STREET: 1455 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/908,597A
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/440,391
; FILING DATE: 12-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: WIXON, HENRY N.
; REGISTRATION NUMBER: 32,073
; REFERENCE/DOCKET NUMBER: 104322.147
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-942-8400
; TELEFAX: 202-942-8484
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 36
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-908-597A-14

Query Match      94.4%; Score 85; DB 1; Length 36;
Best Local Similarity 100.0%; Pred. No. 3.1e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GQVGRQLAIGDDINRR 17
|||||
Db      8 GQVGRQLAIGDDINRR 24
|||||

RESULT 13
US-08-908-597A-14
; Sequence 14, Application US/08908597A
; Patent No. 5863795
; GENERAL INFORMATION:
; APPLICANT: CHITTENDEN, Thomas D.; and
; APPLICANT: LUTZ, Robert J.
; TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH
; MODULATE APOPTOSIS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hale and Dorr
; STREET: 1455 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/908,597A
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/440,391
; FILING DATE: 12-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: WIXON, HENRY N.
; REGISTRATION NUMBER: 32,073
; REFERENCE/DOCKET NUMBER: 104322.147
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-942-8400
; TELEFAX: 202-942-8484
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 36
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-908-597A-14
```

Query Match 94.4%; Score 85; DB 2; Length 36;
Best Local Similarity 100.0%; Pred. No. 3.1e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GQVGRQLAIIGDDINRR 17
DB 8 GQVGRQLAIIGDDINRR 24

RESULT 14
US-09-236-385A-14
; Sequence 14, Application US/09236385A
; Patent No. 6221615
; GENERAL INFORMATION:
; APPLICANT: CHITTENDEN, Thomas D.; and
; LUTZ, Robert J.
; TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH
; MODULATE APOPTOSIS
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hale and Dorr
; STREET: 1455 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; ZIP: 20004

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/236,385A
FILING DATE: 25-Jan-1999
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: WIXON, HENRY N.
REGISTRATION NUMBER: 32,073
(C) ATTORNEY DOCKET NO. 104322.147CIP

TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-942-8400
TELEFAX: 202-942-8484
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 36 base pairs
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-236-385A-14

Query Match 94.4%; Score 85; DB 3; Length 36;
Best Local Similarity 100.0%; Pred. No. 3.1e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GQVGRQLAIIGDDINRR 17
DB 8 GQVGRQLAIIGDDINRR 24

RESULT 15
PCT-US96-06122-14
; Sequence 14, Application PC/TUS9606122
; GENERAL INFORMATION:
; APPLICANT: IMMUNOGEN, INC.
; TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS
; MODULATE APOPTOSIS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hale and Dorr
; STREET: 1455 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.

ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/06122
FILING DATE: HERewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/440,391
FILING DATE: 12-MAY-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: WIXON, HENRY N.
REGISTRATION NUMBER: 32,073
REFERENCE/DOCKET NUMBER: 104322.147PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-942-8400
TELEFAX: 202-942-8484
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 36 base pairs
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US96-06122-14

Query Match 94.4%; Score 85; DB 5; Length 36;
Best Local Similarity 100.0%; Pred. No. 3.1e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GQVGRQLAIIGDDINRR 17
DB 8 GQVGRQLAIIGDDINRR 24

Search completed: November 10, 2004, 13:44:11
Job time : 7.21762 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 16:36:12 ; Search time 19.3368 Seconds
(without alignments)

328.807 Million cell updates/sec

Title: US-10-092-750-41

Perfect score: 90

Sequence: 1 GOVGRQLAIIGDDINERK 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/1/pubaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90	100.0	18	14	US-10-092-750-2
2	90	100.0	18	14	US-10-092-750-41
3	85	94.4	19	11	US-09-828-870-35
4	85	94.4	20	11	US-09-828-870-36
5	85	94.4	28	11	US-09-828-870-2
6	85	94.4	28	11	US-09-828-870-18
7	85	94.4	36	11	US-09-828-870-14
8	85	94.4	117	14	US-10-189-294-4
9	85	94.4	210	14	US-10-101-482-22
10	85	94.4	211	14	US-10-101-482-7
11	85	94.4	211	14	US-10-101-482-9
12	85	94.4	211	14	US-10-101-482-10
13	85	94.4	211	14	US-10-101-482-11

85	94.4	211	14	US-10-189-294-2	Sequence 2, Appli
85	94.4	211	14	US-10-177-293-25	Sequence 25, Appli
85	94.4	211	15	US-10-003-632C-6	Sequence 6, Appli
80	88.9	16	10	US-09-840-085-73	Sequence 73, Appli
80	88.9	16	14	US-10-158-769-3	Sequence 3, Appli
80	88.9	16	14	US-10-059-461-240	Sequence 240, App
80	88.9	16	17	US-10-729-156-9	Sequence 9, Appli
80	88.9	17	14	US-10-092-750-240	Sequence 240, App
80	88.9	19	14	US-10-196-080-2	Sequence 2, Appli
80	88.9	28	14	US-10-092-750-152	Sequence 152, App
79	87.8	31	11	US-09-828-870-3	Sequence 3, Appli
79	87.8	31	11	US-09-828-870-16	Sequence 16, Appli
74	82.2	15	9	US-09-738-396-11	Sequence 11, Appli
74	82.2	15	11	US-09-828-870-10	Sequence 10, Appli
74	82.2	15	11	US-09-828-870-20	Sequence 20, Appli
74	82.2	15	11	US-09-828-870-37	Sequence 37, Appli
69	76.7	15	11	US-09-828-870-38	Sequence 38, Appli
54	60.0	24	9	US-09-682-667-16	Sequence 16, Appli
53	58.9	15	10	US-09-840-085-24	Sequence 24, Appli
50	55.6	226	15	US-10-425-114-56560	Sequence 56560, A
50	55.6	444	17	US-10-425-115-261716	Sequence 261716,
47	52.2	15	9	US-09-913-599-15	Sequence 15, Appli
47	52.2	165	9	US-09-913-599-9	Sequence 9, Appli
47	52.2	195	9	US-09-913-599-8	Sequence 8, Appli
46	51.1	15	10	US-09-840-085-25	Sequence 25, Appli
46	51.1	720	15	US-10-282-122A-72880	Sequence 72880, A
46	51.1	833	15	US-10-282-122A-75311	Sequence 75311, A
46	51.1	834	9	US-09-815-242-10076	Sequence 10076, A
46	51.1	834	14	US-10-369-493-741	Sequence 741, App
46	51.1	834	15	US-10-282-122A-43390	Sequence 43390, A
46	51.1	842	9	US-09-815-242-13718	Sequence 13718, A
46	51.1	915	15	US-10-282-122A-77408	Sequence 77408, A

ALIGNMENTS

RESULT 1

US-10-092-750-2
; Sequence 2, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; APPLICANT: Wright, Martin C.
; TITLE OF INVENTION: Polypeptides Interactive with BCL-X1
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092,750
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/274,526
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-2

Query Match 100.0%; Score 90; DB 14; Length 18;
Best Local Similarity 100.0%; Pred. No. 5.7e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GOVGRQLAIIGDDINERK 18

DB 1 GOVGRQLAIIGDDINERK 18

RESULT 2

US-10-092-750-41
; Sequence 41, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:

APPLICANT: Hammond, Philip W.
APPLICANT: Alpin, Julia
APPLICANT: Wright, Martin C.
TITLE OF INVENTION: Polypeptides Interactive with BCL-XL
FILE REFERENCE: 50036/050002
CURRENT APPLICATION NUMBER: US/10/092,750
CURRENT FILING DATE: 2002-03-07
PRIOR APPLICATION NUMBER: US 60/274,526
PRIOR FILING DATE: 2001-03-08
NUMBER OF SEQ ID NOS: 253
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 41
LENGTH: 18
TYPE: PRT
ORGANISM: Homo sapiens
US-10-092-750-41

Query Match 100.0%; Score 90; DB 14; Length 18;
Best Local Similarity 100.0%; Pred. No. 5.7e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GQVGRQLAIIGDDINRRK 18
| | | | | | | | | | | | | | | | | |
DB 1 GQVGRQLAIIGDDINRRK 18

RESULT 3
US-09-828-870-35
Sequence 35, Application US/09828870
Publication No. US20040054129A1
GENERAL INFORMATION:
APPLICANT: CHITTENDEN, Thomas D.; and
LUTZ, Robert J.
TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH
MODULATE APOPTOSIS

NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hale and Dorr
STREET: 1455 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
ZIP: 20004

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/828,870
FILING DATE: 10-Apr-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/236,385
FILING DATE: 25-JANUARY-1999
ATTORNEY/AGENT INFORMATION:
NAME: WIXON, HENRY N.
REGISTRATION NUMBER: 32,073
(C) ATTORNEY DOCKET NO. 104322.147CIP

TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-942-8400
TELEFAX: 202-942-8484

INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 19 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: Peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 35:
US-09-828-870-35

Query Match 94.4%; Score 85; DB 11; Length 19;
Best Local Similarity 100.0%; Pred. No. 4.1e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GQVGRQLAIIGDDINRR 17
| | | | | | | | | | | | | | | | | |
DB 2 GQVGRQLAIIGDDINRR 18

RESULT 4
US-09-828-870-36
Sequence 36, Application US/09828870
Publication No. US20040054129A1
GENERAL INFORMATION:
APPLICANT: CHITTENDEN, Thomas D.; and
LUTZ, Robert J.
TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH
MODULATE APOPTOSIS
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hale and Dorr
STREET: 1455 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
ZIP: 20004

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/828,870
FILING DATE: 10-Apr-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/236,385
FILING DATE: 25-JANUARY-1999

ATTORNEY/AGENT INFORMATION:
NAME: WIXON, HENRY N.
REGISTRATION NUMBER: 32,073
(C) ATTORNEY DOCKET NO. 104322.147CIP

TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-942-8400
TELEFAX: 202-942-8484
INFORMATION FOR SEQ ID NO: 36
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 36

US-09-828-870-36
Query Match 94.4%; Score 85; DB 11; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.4e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GQVGRQLAIIGDDINRR 17
| | | | | | | | | | | | | | | | | |
DB 3 GQVGRQLAIIGDDINRR 19

RESULT 5
US-09-828-870-2
Sequence 2, Application US/09828870
Publication No. US20040054129A1
GENERAL INFORMATION:
APPLICANT: CHITTENDEN, Thomas D.; and
LUTZ, Robert J.
TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH
MODULATE APOPTOSIS
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hale and Dorr
STREET: 1455 Pennsylvania Avenue, N.W.
CITY: Washington

STATE: D.C.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/828,870
FILING DATE: 10-Apr-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/236,385
FILING DATE: 25-JANUARY-1999
ATTORNEY/AGENT INFORMATION:
NAME: WIXON, HENRY N.
REGISTRATION NUMBER: 32,073
(C) ATTORNEY DOCKET NO. 104322.147CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-942-8400
TELEFAX: 202-942-8484
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 amino acid
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-828-870-2
Query Match 94.4%; Score 85; DB 11; Length 28;
Best Local Similarity 100.0%; Pred. No. 6.3e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GQVGRQLAIIGDDINRR 17
Db 6 GQVGRQLAIIGDDINRR 22
RESULT 6
US-09-828-870-18
Sequence 18, Application US/09828870
Publication No. US20040054129A1
GENERAL INFORMATION:
APPLICANT: CHITTENDEN, Thomas D.; and
LUTZ, Robert J.
TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH
MODULATE APOPTOSIS
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hale and Dorr
STREET: 1455 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/828,870
FILING DATE: 10-Apr-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/236,385
FILING DATE: 25-JANUARY-1999
ATTORNEY/AGENT INFORMATION:
NAME: WIXON, HENRY N.
REGISTRATION NUMBER: 32,073
(C) ATTORNEY DOCKET NO. 104322.147CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-942-8400

TELEFAX: 202-942-8484
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-09-828-870-18
Query Match 94.4%; Score 85; DB 11; Length 28;
Best Local Similarity 100.0%; Pred. No. 6.3e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GQVGRQLAIIGDDINRR 17
Db 6 GQVGRQLAIIGDDINRR 22
RESULT 7
US-09-828-870-14
Sequence 14, Application US/09828870
Publication No. US20040054129A1
GENERAL INFORMATION:
APPLICANT: CHITTENDEN, Thomas D.; and
LUTZ, Robert J.
TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH
MODULATE APOPTOSIS
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hale and Dorr
STREET: 1455 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/828,870
FILING DATE: 10-Apr-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/236,385
FILING DATE: 25-JANUARY-1999
ATTORNEY/AGENT INFORMATION:
NAME: WIXON, HENRY N.
REGISTRATION NUMBER: 32,073
(C) ATTORNEY DOCKET NO. 104322.147CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-942-8400
TELEFAX: 202-942-8484
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 36 base pairs
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-828-870-14
Query Match 94.4%; Score 85; DB 11; Length 36;
Best Local Similarity 100.0%; Pred. No. 8.3e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GQVGRQLAIIGDDINRR 17
Db 8 GQVGRQLAIIGDDINRR 24
RESULT 8

US-10-189-294-4
; Sequence 0, Application US/10189294
; Publication No. US20030060615A1
; GENERAL INFORMATION:
; APPLICANT: Kiefer, Michael C.
; Gibson, Helen L.
; Fitzpatrick, Paul A.
; Barr, Philip J.
; TITLE OF INVENTION: A NOVEL BAK BINDING PROTEIN, DNA
; ENCODING THE PROTEIN, AND METHODS OF USE THEREOF
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWN, THERESA A.
; STREET: 1560 Broadway, Suite 1200
; CITY: Denver
; STATE: CO
; COUNTRY: USA
; ZIP: 80202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/189,294
; FILING DATE: 01-Jul-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/381,488
; FILING DATE: 11-Feb-2000
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWN, THERESA A.
; REGISTRATION NUMBER: 32,547
; REFERENCE/DOCKET NUMBER: 4147-15-PUS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: /note= "Bak (delta)2 (delta)TM"
; SEQUENCE CHARACTERISTICS:
; LENGTH: 117 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..117
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-189-294-4
Query Match 94.4%; Score 85; DB 14; Length 117;
Best Local Similarity 100.0%; Pred. No. 3e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GOVGRQLAIIGDDINRR 17
Db 2 GOVGRQLAIIGDDINRR 18
RESULT 9
US-10-101-482-22
; Sequence 22, Application US/10101482
; Publication No. US2003008837A1
; GENERAL INFORMATION:
; APPLICANT: KIEFER, MICHAEL C.
; BARR, PHILIP J.
; TITLE OF INVENTION: NOVEL APOPTOSIS-MODULATING PROTEINS, DNA
; ENCODING THE PROTEINS AND METHODS OF USE THEREOF
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/101,482
; FILING DATE: 18-Mar-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/320,157
; FILING DATE: 07-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: LEHNHARDT, SUSAN K.
; REGISTRATION NUMBER: 33,943
; REFERENCE/DOCKET NUMBER: 23647-20007.20

COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/101,482
FILING DATE: 18-Mar-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/320,157
FILING DATE: 07-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: LEHNHARDT, SUSAN K.
REGISTRATION NUMBER: 33,943
REFERENCE/DOCKET NUMBER: 23647-20007.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 813-5600
TELEFAX: (415) 494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 210 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 22:
US-10-101-482-22
Query Match 94.4%; Score 85; DB 14; Length 210;
Best Local Similarity 100.0%; Pred. No. 5.6e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GOVGRQLAIIGDDINRR 17
Db 72 GOVGRQLAIIGDDINRR 88
RESULT 10
US-10-101-482-7
; Sequence 7, Application US/10101482
; Publication No. US2003008837A1
; GENERAL INFORMATION:
; APPLICANT: KIEFER, MICHAEL C.
; BARR, PHILIP J.
; TITLE OF INVENTION: NOVEL APOPTOSIS-MODULATING PROTEINS, DNA
; ENCODING THE PROTEINS AND METHODS OF USE THEREOF
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/101,482
; FILING DATE: 18-Mar-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/320,157
; FILING DATE: 07-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: LEHNHARDT, SUSAN K.
; REGISTRATION NUMBER: 33,943
; REFERENCE/DOCKET NUMBER: 23647-20007.20

TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 813-5600
 TELEFAX: (415) 494-0792
 TELEX: 706141
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 211 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 7:
 US-10-101-482-7

Query Match 94.4%; Score 85; DB 14; Length 211;
 Best Local Similarity 100.0%; Pred. No. 5.6e-06;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GQVGRQLAIGDDINRR 17
 |||||
 Db 72 GQVGRQLAIGDDINRR 88

RESULT 11

US-10-101-482-9
 ; Sequence 9, Application US/10101482
 ; Publication No. US2003000837A1
 ; GENERAL INFORMATION:
 ; APPLICANT: KIEFER, MICHAEL C.
 ; BARR, PHILIP J.

TITLE OF INVENTION: NOVEL APOPTOSIS-MODULATING PROTEINS, DNA
 ENCODING THE PROTEINS AND METHODS OF USE THEREOF

NUMBER OF SEQUENCES: 22
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: MORRISON & FOERSTER
 STREET: 755 Page Mill Road
 CITY: Palo Alto
 STATE: California
 COUNTRY: USA
 ZIP: 94304-1018

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/101,482
 FILING DATE: 18-Mar-2002
 CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/320,157
 FILING DATE: 07-OCT-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: LEHNHARDT, SUSAN K.
 REGISTRATION NUMBER: 33,943
 REFERENCE/DOCKET NUMBER: 23647-20007.20
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 813-5600
 TELEFAX: (415) 494-0792
 TELEX: 706141

INFORMATION FOR SEQ ID NO: 9:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 211 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 9:
 US-10-101-482-9

Query Match 94.4%; Score 85; DB 14; Length 211;
 Best Local Similarity 100.0%; Pred. No. 5.6e-06;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GQVGRQLAIGDDINRR 17

Db 72 GQVGRQLAIGDDINRR 88
 |||||

RESULT 12

US-10-101-482-10
 ; Sequence 10, Application US/10101482
 ; Publication No. US2003000837A1
 ; GENERAL INFORMATION:
 ; APPLICANT: KIEFER, MICHAEL C.
 ; BARR, PHILIP J.

TITLE OF INVENTION: NOVEL APOPTOSIS-MODULATING PROTEINS, DNA
 ENCODING THE PROTEINS AND METHODS OF USE THEREOF

NUMBER OF SEQUENCES: 22
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: MORRISON & FOERSTER
 STREET: 755 Page Mill Road
 CITY: Palo Alto
 STATE: California
 COUNTRY: USA
 ZIP: 94304-1018

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/101,482
 FILING DATE: 18-Mar-2002
 CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/320,157
 FILING DATE: 07-OCT-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: LEHNHARDT, SUSAN K.
 REGISTRATION NUMBER: 33,943
 REFERENCE/DOCKET NUMBER: 23647-20007.20
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 813-5600
 TELEFAX: (415) 494-0792
 TELEX: 706141

INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 211 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 SEQUENCE DESCRIPTION: SEQ ID NO: 10:
 US-10-101-482-10

Query Match 94.4%; Score 85; DB 14; Length 211;
 Best Local Similarity 100.0%; Pred. No. 5.6e-06;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GQVGRQLAIGDDINRR 17
 |||||
 Db 72 GQVGRQLAIGDDINRR 88

RESULT 13

US-10-101-482-11
 ; Sequence 11, Application US/10101482
 ; Publication No. US2003000837A1
 ; GENERAL INFORMATION:
 ; APPLICANT: KIEFER, MICHAEL C.
 ; BARR, PHILIP J.

TITLE OF INVENTION: NOVEL APOPTOSIS-MODULATING PROTEINS, DNA
 ENCODING THE PROTEINS AND METHODS OF USE THEREOF

NUMBER OF SEQUENCES: 22
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: MORRISON & FOERSTER
 STREET: 755 Page Mill Road
 CITY: Palo Alto

STATE: California
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/101,482
FILING DATE: 18-Mar-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/330,157
FILING DATE: 07-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: LEHNHARDT, SUSAN K.
REGISTRATION NUMBER: 33,943
REFERENCE/DOCKET NUMBER: 23647-20007.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 813-5600
TELEFAX: (415) 494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 211 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-10-101-482-11
Query Match 94.4%; Score 85; DB 14; Length 211;
Best Local Similarity 100.0%; Pred. No. 5.6e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GOVGRQLAIIIGDDINRR 17
DB 72 GOVGRQLAIIIGDDINRR 88
RESULT 14
US-10-189-294-2
Sequence 2, Application US/10189294
Publication No. US20030080615A1
GENERAL INFORMATION:
APPLICANT: Kiefer, Michael C.
Gibson, Helen L.
Fitzpatrick, Paul A.
Barr, Philip J.
TITLE OF INVENTION: A NOVEL BAK BINDING PROTEIN, DNA
ENCODING THE PROTEIN, AND METHODS OF USE THEREOF
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWN, THERESA A.
STREET: 1560 Broadway, Suite 1200
CITY: Denver
STATE: CO
COUNTRY: USA
ZIP: 80202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/189,294
FILING DATE: 01-Jul-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/381,488
FILING DATE: 11-Feb-2000
ATTORNEY/AGENT INFORMATION:

NAME: BROWN, THERESA A.
REGISTRATION NUMBER: 32,547
REFERENCE/DOCKET NUMBER: 4147-15-PUS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 211 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-189-294-2
Query Match 94.4%; Score 85; DB 14; Length 211;
Best Local Similarity 100.0%; Pred. No. 5.6e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GOVGRQLAIIIGDDINRR 17
DB 72 GOVGRQLAIIIGDDINRR 88
RESULT 15
US-10-177-293-25
Sequence 25, Application US/10177293
Publication No. US20030124128A1
GENERAL INFORMATION:
APPLICANT: Lillie, James
Glat, Karen
APPLICANT: Zhao, Xumei
APPLICANT: Gannavarpu, Manjula
APPLICANT: Kamatkar, Shubhangi
APPLICANT: Mertens, Maureen
APPLICANT: Myer, Vic
APPLICANT: Wang, Youzhen
APPLICANT: Xu, Yongyao
APPLICANT: Hoersch, Sebastian
APPLICANT: Monahan, John
APPLICANT: Meyers, Rachel E.
APPLICANT: East Jr., Robert C.
APPLICANT: Hortobagyi, Gabriel M.
APPLICANT: Pusztai, Lajos
APPLICANT: Meric, Funda
APPLICANT: Sahin, Aysegul
APPLICANT: Mills, Gordon B.
TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
PREVENTION, AND THERAPY OF BREAST CANCER
FILE REFERENCE: MRI-038
CURRENT APPLICATION NUMBER: US/10/177,293
CURRENT FILING DATE: 2002-06-21
PRIOR APPLICATION NUMBER: US 60/299,887
PRIOR FILING DATE: 2001-06-21
PRIOR APPLICATION NUMBER: US 60/301,572
PRIOR FILING DATE: 2001-06-27
PRIOR APPLICATION NUMBER: US 60/306,501
PRIOR FILING DATE: 2001-07-18
PRIOR APPLICATION NUMBER: US 60/325,002
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: US 60/362,585
PRIOR FILING DATE: 2002-03-05
PRIOR APPLICATION NUMBER: US 60/xxx,xxx
PRIOR FILING DATE: 2002-05-14
NUMBER OF SEQ ID NOS: 506
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 25
LENGTH: 211
TYPE: PRT
ORGANISM: Homo sapiens
US-10-177-293-25
Query Match 94.4%; Score 85; DB 14; Length 211;

Best Local Similarity 100.0%; Pred. No. 5.6e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GQVGRQLAIIGDDINRR 17
Db 72 GQVGRQLAIIGDDINRR 88

Search completed: November 11, 2004, 01:28:23
Job time : 20.3868 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 12:29:32 ; Search time 4.01036 Seconds
(without alignments)
431.857 Million cell updates/sec

Title: US-10-092-750-41
Perfect score: 90
Sequence: 1 GQVGRQLAIIGDDINRRK 18
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: Pir1.*
2: Pir2.*
3: Pir3.*
4: Pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	85	94.4	211	2 S58873	Bak protein - huma
2	85	94.4	211	2 S58875	cdn-2 protein - hu
3	46.5	51.7	357	2 T06308	protein phosphatas
4	46	51.1	833	2 AE0564	H+/K+-exchanging A
5	46	51.1	834	1 C54779	probable copper-tr
6	46	51.1	834	2 E85546	probable ATPase yb
7	46	51.1	834	2 A90696	Cu(II)-translocatio
8	46	51.1	915	2 H82104	cation transport A
9	46	51.1	1226	2 S44824	F54F2.1 protein -
10	45	50.0	426	2 S58684	phosphopyruvate hy
11	45	50.0	426	2 H71967	enolase - Helicoba
12	44.5	49.4	532	2 JN0084	phytoene dehydroge
13	44	48.9	234	2 AC3205	IS66 family Orf4 (
14	44	48.9	258	2 H75027	sv v-atpase proteo
15	44	48.9	261	2 B71213	probable chemorece
16	44	48.9	475	2 T48031	hypothetical prote
17	44	48.9	593	2 S75352	ABC-type transport
18	44	48.9	693	2 G82618	plus biogenesis p
19	44	48.9	803	1 E70041	probable copper-tr
20	43.5	48.3	355	2 G84643	probable protein p
21	43	47.8	381	2 G71906	probable transcrip
22	43	47.8	381	2 G64607	response regulator
23	43	47.8	444	2 TQ1185	phosphopyruvate hy
24	43	47.8	446	2 T03257	probable phosphopy
25	43	47.8	446	2 T02221	phosphopyruvate hy
26	43	47.8	447	2 G86940	probable enolase [
27	43	47.8	455	2 E90316	oxidoreductase [im
28	43	47.8	664	2 D96633	hypothetical prote
29	43	47.8	770	2 T23999	hypothetical prote

30	43	47.8	827	2 B95969	probable H+/K+-exc
31	42	46.7	356	2 S71460	ribose-phosphate d
32	42	46.7	356	2 A53433	ribose-phosphate d
33	42	46.7	482	2 S60757	catalase (EC 1.11.
34	42	46.7	637	2 D97708	cell division prot
35	42	46.7	637	2 C71712	cell division prot
36	42	46.7	826	2 D95330	cu2+-exporting ATP
37	41	45.6	70	2 H71313	hypothetical prote
38	41	45.6	173	2 A85799	crossover junction
39	41	45.6	173	2 D38113	crossover junction
40	41	45.6	173	2 E90950	Holliday junction
41	41	45.6	251	2 T44678	chemotaxis protein
42	41	45.6	329	2 S67451	hypothetical prote
43	41	45.6	447	2 T13091	probable minor cap
44	41	45.6	530	2 C72291	methyl-accepting c
45	41	45.6	539	2 F72288	methyl-accepting c

ALIGNMENTS

RESULT 1

S58873
Bak protein - human
N;Alternate names: bcl-2 homolog; cdn-1 protein
C;Species: Homo sapiens (man)
C;Date: 15-Feb-1996 #sequence revision 01-Mar-1996 #text_change 09-Jul-2004
C;Accession: S58873; S58872; S58874
R;Chittenden, T.; Harrington, E.A.; O'Connor, R.; Flemington, C.; Lutz, R.J.; Evan, G.I.;
Nature 374, 733-736, 1995
A;Title: Induction of apoptosis by the Bcl-2 homologue Bak.
A;Reference number: S58873; MUID:95231653; PMID:7715730
A;Accession: S58873
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-211 <CHI>
A;Cross-references: UNIPROT:Q16611; EMBL:U23765; NID:g758797; PIDN:AAA93066.1; PID:g7587;
R;Farrow, S.N.; White, J.H.M.; Martinou, I.; Raven, T.; Pun, K.T.; Grinham, C.J.; Martinou;
Nature 374, 736-739, 1995
A;Title: Cloning of a bcl-2 homologue by interaction with adenovirus ElB 19K.
A;Reference number: S58872; MUID:95231652; PMID:7715729
A;Accession: S58872
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-211 <FAR>
A;Cross-references: EMBL:X84213; NID:G804984; PIDN:CAA58997.1; PID:G804985
R;Kiefer, M.C.; Brauer, M.J.; Powers, V.C.; Wu, J.J.; Unanue, S.R.; Tomei, L.D.; Barr, J;
Nature 374, 736-739, 1995
A;Title: Modulation of apoptosis by the widely distributed Bcl-2 homologue Bak.
A;Reference number: S58874; MUID:95231654; PMID:7715731
A;Accession: S58874
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-211 <KIE>
A;Cross-references: EMBL:U16811; NID:g595923; PIDN:AAA74466.1; PID:g595924
C;Genetics:
A;Gene: GDB:BAK
A;Cross-references: GDB:635987

Query Match 94.4%; Score 85; DB 2; Length 211;
Best Local Similarity 100.0%; Pred.No. 1.2e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GQVGRQLAIIGDDINRR 17

Db 72 GQVGRQLAIIGDDINRR 88

RESULT 2

S58875
cdn-2 protein - human
C;Species: Homo sapiens (man)
C;Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 09-Jul-2004

C;Accession: S58875
 R;Kiefer, M.C.; Brauer, M.J.; Powers, V.C.; Wu, J.J.; Umansky, S.R.; Tomei, L.D.; Barr, Nature 374, 736-739, 1995
 A;Title: Modulation of apoptosis by the widely distributed Bcl-2 homologue Bak.
 A;Reference number: S58874; MUID:95231654; PMID:7715731
 A;Accession: S58875
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-211 <KIE>
 A;Cross-references: UNIPROT:Q13014; EMBL:U16812; NID:Q595925; PIDN:AAA74467.1; PID:Q5959
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1994

Query Match 94.4%; Score 85; DB 2; Length 211;
 Best Local Similarity 100.0%; Pred. No. 1.2e-06;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGVGQLAIIGDINRR 17
 |||||
 Db 72 GGVGQLAIIGDINRR 88

RESULT 3

T06308
 protein phosphatase 2C homolog Fl1C18.60 - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 09-Jul-2004
 C;Accession: T06308
 R;Bevan, M.; Terry, K.F.X.; Schueller, C.
 ewes, H.W.; Mayer, K.F.X.; Schueller, C.
 submitted to the Protein Sequence Database, April 1999
 A;Reference number: Z15589
 A;Accession: T06308
 A;Molecule type: DNA
 A;Residues: 1-357 <BEV>
 A;Cross-references: UNIPROT:Q9S253; EMBL:AL049607; GSPDB:GN00062; ATSP:FL1C18.60
 A;Experimental source: cultivar Columbia; BAC clone Fl1C18
 C;Genetics:
 A;Gene: ATSP:Fl1C18.60
 A;Map position: 4
 A;Introns: 3/3; 61/1; 97/2; 148/3; 190/3; 232/1; 257/3; 275/2; 293/3
 C;Superfamily: human phosphoprotein phosphatase 1A

Query Match 51.7%; Score 46.5; DB 2; Length 357;
 Best Local Similarity 58.8%; Pred. No. 6;
 Matches 10; Conservative 4; Mismatches 2; Indels 1; Gaps 1;

QY 1 GGVGQLAIIGDINRR 16
 |||||
 Db 104 GGVGQLAIIGDINRR 120

RESULT 4

H564
 H+/K+-exchanging ATPase (EC 3.6.3.10) - Salmonella enterica subsp. enterica serovar Typh
 C;Species: Salmonella enterica subsp. enterica serovar Typhi
 A;Note: this species has also been called Salmonella typhi
 C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 C;Accession: AE0564
 R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
 th, T.; Conerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
 S.; Moule, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A;Authors: Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
 A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
 A;Reference number: AB0502; MUID:21534947; PMID:11677608
 A;Accession: AE0564
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-833 <PAR>
 A;Cross-references: GB:AL513382; PIDN:CAD04983.1; PID:gl6501768; GSPDB:GN00176
 C;Genetics:
 A;Gene: STY0544
 C;Superfamily: Bacillus probable copper-transporting ATPase yvgX; ATPase nucleotide-bind

C;Keywords: hydrolase

Query Match 51.1%; Score 46; DB 2; Length 833;
 Best Local Similarity 66.7%; Pred. No. 19;
 Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 4 GRQLAIIGDDIN 15
 |||||
 Db 711 GRQVAVGDDIN 722

RESULT 5

C64779
 probable copper-transporting ATPase (EC 3.6.1.-) - Escherichia coli (strain K-12)
 C;Species: Escherichia coli
 C;Date: 05-Dec-1998 #sequence_revision 05-Dec-1998 #text_change 09-Jul-2004
 C;Accession: C64779
 R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Col

A.; Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997
 A;Title: The complete genome sequence of Escherichia coli K-12.
 A;Reference number: A64720; MUID:97426617; PMID:9278503

A;Accession: C64779
 A;Status: nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA

A;Residues: 1-834 <BLAT>
 A;Cross-references: UNIPROT:Q59385; GB:AE000154; GB:U00096; NID:gl786683; PIDN:AACT3586.1
 A;Experimental source: strain K-12, substrain MG1655
 C;Genetics:

C;Gene: ybaR
 C;Superfamily: Bacillus probable copper-transporting ATPase yvgX; ATPase nucleotide-bind
 C;Keywords: ATP; copper binding; hydrolase; ion transport; metal binding; phosphoprotein;
 F;9-38/Domain: heavy-metal-associated homology <HMA1>
 F;105-134/Domain: heavy-metal-associated homology <HMA2>
 F;189-205/Domain: transmembrane #status predicted <TM1>
 F;218-234/Domain: transmembrane #status predicted <TM2>
 F;224-568/Domain: ATPase transduction domain homology <ATT>
 F;438-454/Domain: transmembrane #status predicted <TM3>
 F;468-484/Domain: transmembrane #status predicted <TM4>
 F;631-647/Domain: transmembrane #status predicted <TM5>
 F;643-785/Domain: ATPase nucleotide-binding domain homology <ATN>
 F;806-822/Domain: transmembrane #status predicted <TM6>
 F;108,110,113/Binding site: copper (Met, Cys, Cys) #status predicted
 F;523/Active site: Asp (aspartylphosphate intermediate) #status predicted

Query Match 51.1%; Score 46; DB 1; Length 834;
 Best Local Similarity 66.7%; Pred. No. 19;
 Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 4 GRQLAIIGDDIN 15
 |||||
 Db 712 GRQVAVGDDIN 723

RESULT 6

E85546
 probable ATPase ybaR [imported] - Escherichia coli (strain O157:H7, substrain EDL933)
 C;Species: Escherichia coli
 C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
 C;Accession: E85546

R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew,
 iller, L.; Grotbeck, E.J.; Davis, N.W.; Llin, A.; Dimalanta, E.; Potamoculis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A;Reference number: A85480; MUID:21074935; PMID:11206551

A;Accession: E85546
 A;Status: preliminary
 A;Molecule type: DNA

A;Residues: 1-834 <STO>
 A;Cross-references: UNIPROT:Q8XD24; GB:AE005174; NID:gl12513357; PIDN:AAQ54833.1; GSPDB:GN
 A;Experimental source: strain O157:H7, substrain EDL933
 C;Genetics:
 A;Gene: ybaR

C:Superfamily: Bacillus probable copper-transporting ATPase yvqX; ATPase nucleotide-binding

Query Match 51.1%; Score 46; DB 2; Length 834;
 Best Local Similarity 66.7%; Pred. No. 19;
 Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 4 GRQLAIIIGDDIN 15
 |||:|:|:|
 DB 712 GRQVAMVGDGIN 723

RESULT 7
 A90696
 Cu(II)-translocation P-type ATPase [imported] - Escherichia coli (strain O157:H7, substrate)
 C:Species: Escherichia coli
 C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
 C:Accession: A90696
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.; Sasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 DNA Res. 8, 11-22, 2001
 A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genome
 A:Reference number: A99629; MUID:21156231; PMID:11258796
 A:Accession: A90696
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-834 <HAY>
 A:Cross-references: UNIPROT:Q8XD24; GB:BA000007; PIDN:BA833960.1; PID:g13359994; GSPDB:G
 A:Experimental source: strain O157:H7, substrain RIMD 0509952
 C:Genetics:
 A:Gene: EC0537
 C:Superfamily: Bacillus probable copper-transporting ATPase yvqX; ATPase nucleotide-binding

Query Match 51.1%; Score 46; DB 2; Length 834;
 Best Local Similarity 66.7%; Pred. No. 19;
 Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 4 GRQLAIIIGDDIN 15
 |||:|:|:|
 DB 712 GRQVAMVGDGIN 723

RESULT 8
 H82104
 cation transport ATPase, El-E2 family VC2215 [imported] - Vibrio cholerae (strain N16961)
 C:Species: Vibrio cholerae
 C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
 C:Accession: H82104
 R:Heidelberg, J.F.; Risen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.; chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, H.
 l. R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 Nature 406, 477-483, 2000
 A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
 A:Reference number: A82035; MUID:20406833; PMID:10952301
 A:Accession: H82104
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-315 <HEI>
 A:Cross-references: UNIPROT:Q9KPD7; GB:AB004293; GB:AB003852; NID:g9656766; PIDN:AAF9535
 A:Experimental source: serogroup O1; strain N16961; biotype El Tor
 C:Genetics:
 A:Gene: VC2215
 A:Map position: 1
 C:Superfamily: Bacillus probable copper-transporting ATPase yvqX; ATPase nucleotide-binding

Query Match 51.1%; Score 46; DB 2; Length 915;
 Best Local Similarity 64.3%; Pred. No. 21;
 Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 QVGRQLAIIIGDDIN 15
 |||:|:|:|
 DB 786 QQRKXVAMVGDGIN 799

RESULT 9
 S44824
 F54r21.1 protein - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 14-Sep-1994 #sequence_revision 12-May-1995 #text_change 09-Jul-2004
 C:Accession: S44824
 R:Anderson, K.
 submitted to the EMBL Data Library, September 1993
 A:Description: Sequence of the C. elegans cosmid F54r2.
 A:Reference number: S44817
 A:Accession: S44824
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1226 <AND>
 A:Cross-references: UNIPROT:P34446; EMBL:L23645; NID:g388603; PID:g388605
 C:Genetics:
 A:Introns: 58/2; 137/3; 179/1; 316/2; 393/1; 551/3; 597/2; 662/2; 899/3; 1178/3
 C:Keywords: cytoskeleton; transmembrane protein

Query Match 51.1%; Score 46; DB 2; Length 1226;
 Best Local Similarity 53.8%; Pred. No. 29;
 Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 QVGRQLAIIIGDD 13
 |||:|:|:|
 DB 359 GVFGKQIAVVGDD 371

RESULT 10
 S58684
 phosphopyruvate hydratase (EC 4.2.1.11) - Helicobacter pylori (strains 26695 and others)
 N:Alternate names: enolase
 C:Species: Helicobacter pylori
 C>Date: 29-Nov-1995 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
 C:Accession: B64539; S58684
 R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khairak, H.G.; Glodek, A.; McKenney, son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.
 Nature 388, 539-547, 1997
 A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.I.
 A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
 A:Reference number: A64520; MUID:97394467; PMID:9252185
 A:Accession: B64539
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-426 <TOM>
 A:Cross-references: UNIPROT:P48285; GB:AE000536; GB:AE000511; NID:g2313230; PIDN:AA07215
 A:Experimental source: strain 26695
 R:Schmitt, W.; Odenbreit, S.; Heuermann, D.; Haas, R.
 Mol. Gen. Genet. 249, 563-572, 1995
 A:Title: Cloning of the Helicobacter pylori rcsA gene and functional characterization of
 A:Reference number: S58683; MUID:96027928; PMID:7476956
 A:Accession: S58684
 A:Molecule type: DNA
 A:Residues: 1-25, '1', 27-68 <SCH>
 A:Cross-references: EMBL:Z35478
 C:Genetics:
 A:Gene: HP0154
 C:Function:
 A:Description: catalyzes the reversible dehydration of 2-phospho-D-glyceric acid to phospho-D-glycerate
 A:Pathway: glycolysis
 C:Superfamily: enolase
 C:Keywords: carbon-oxygen lyase; gluconeogenesis; glycolysis; hydro-lyase; magnesium
 F:42/Binding site: magnesium 2 (Ser) #status predicted
 F:205.338/Active site: Glu, Lys #status predicted
 F:242.286,313/Binding site: magnesium 1 (Asp, Glu, Asp) #status predicted

Query Match 50.0%; Score 45; DB 2; Length 426;
 Best Local Similarity 46.2%; Pred. No. 13;
 Matches 6; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

QY 2 QVGRQLAIIIGDDI 14
 |||:|:|:|

```

Db      303 ELGROIQLVGDDL 315

RESULT 11
H71967
enclase - Helicobacter pylori (strain J99)
C:Species: Helicobacter pylori
A:Variety: strain J99
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
C:Accession: H71967
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;
Nature 397, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
A:Reference number: A71800; MUID:99120557; PMID:9923682
A:Accession: H71967
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-426 <ARN>
A:Cross-references: UNIPROT:Q9ZMS6; GB:AE001453; GB:AE001439; NID:g4154651; PIDN:AAD0572
A:Experimental source: strain J99
C:Genetics:
A:Gene: eno
C:Superfamily: enolase

Query Match      50.0%; Score 45; DB 2; Length 426;
Best Local Similarity 46.2%; Pred. No. 13;
Matches 6; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

QY      2 QVGRQLAIIGDDI 14
Db      303 ELGROIQLVGDDL 315

RESULT 12
JN0084
phytoene dehydrogenase (EC 1.3.1.1) - Aphanocapsa sp. (PCC 6714)
N:Alternate names: phytoene desaturase
C:Species: Aphanocapsa sp.
C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 09-Jul-2004
C:Accession: JN0084
R:Schmidt, A.; Sandmann, G.
Gene 91, 113-117, 1990
A:Title: Cloning and nucleotide sequence of the crtI gene encoding phytoene dehydrogenase
A:Reference number: JN0084; MUID:90382685; PMID:2119326
A:Molecule type: DNA
A:Residues: 1-532 <SCH>
A:Cross-references: UNIPROT:P21134; GB:M55647; GB:M33237; NID:g141989; PIDN:AAA62573.1;
A:Note: the authors translated the codon CAG for residue 380 as Gly
C:Comment: This enzyme catalyzes the symmetrical introduction of two double bonds at C-1
C:Genetics:
A:Gene: crtI
C:Superfamily: Aphanocapsa phytoene dehydrogenase
C:Keywords: carotenoid biosynthesis; membrane bound; oxidoreductase

Query Match      49.4%; Score 44.5; DB 2; Length 532;
Best Local Similarity 45.0%; Pred. No. 20;
Matches 9; Conservative 6; Mismatches 2; Indels 3; Gaps 1;

QY      1 GOVGRQLAI---IGDDINER 17
Db      141 GQTRRLQLLEPIGEDVHRQ 160

RESULT 13
AC3205
IS66 family Orf4 (truncated) [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C:Accession: AC3205
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.;
Erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, I.
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AC3205
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-234 <KUR>
A:Cross-references: UNIPROT:Q8UUV6; GB:AE008687; PIDN:AAL46057.1; PID:g17743817; GSPDB:G
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Orf4*
A:Genome: plasmid

Query Match      48.9%; Score 44; DB 2; Length 234;
Best Local Similarity 46.7%; Pred. No. 9;
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY      3 VGRQLAIIGDDINRR 17
Db      12 VGRQKVLIGDVSR 26

RESULT 14
H75027
sy v-atpase proteolipid PAB1189 - Pyrococcus abyssi (strain Orsay)
C:Species: Pyrococcus abyssi
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C:Accession: H75027
R:anonymous, Genoscope
A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome struc
A:Reference number: A75001
A:Accession: H75027
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-258 <KAW>
A:Cross-references: UNIPROT:Q9UXV1; GB:AJ248288; GB:AL096836; NID:gs458960; PIDN:CAB50662;
A:Experimental source: strain Orsay
C:Genetics:
A:Gene: PAB1189

Query Match      48.9%; Score 44; DB 2; Length 258;
Best Local Similarity 43.8%; Pred. No. 11;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY      1 GOVGRQLAIIGDDINR 16
Db      122 GEAGRGFAVVADEIRR 137

RESULT 15
B71213
probable chemoreceptor protein - Pyrococcus horikoshii
C:Species: Pyrococcus horikoshii
C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 09-Jul-2004
C:Accession: B71213
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekine
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi,
DNA Res. 5, 55-76, 1998
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a
A:Reference number: A71000; MUID:98344137; PMID:9679194
A:Accession: B71213
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-261 <KAW>
A:Cross-references: UNIPROT:O57733; GB:AP000007; NID:g3236134; PIDN:BAA31097.1; PID:g3259
A:Experimental source: strain OT3
A:Note: this accession replaces an interim accession for a sequence replaced by GenBank
C:Genetics:
A:Gene: PH1970

```

Query Match 48.9%; Score 44; DB 2; Length 261;
 Best Local Similarity 43.8%; Pred. No. 11;
 Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 GQVGRQLAIIIGDDINR 16
 |::|||::|::|
 DB 125 GEAGRGFAVVADEIRR 140

Search completed: November 10, 2004, 13:40:49
 Job time : 5.01036 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 12:27:34 ; Search time 21.9793 Seconds
(without alignments)
471.205 Million cell updates/sec

Title: US-10-092-750-41
Perfect score: 90
Sequence: 1 GQVGRQLAIGDDINRRK 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot_02:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	85	94.4	80	2 O77738	O77738 sus scrofa
2	85	94.4	190	2 Q8NFF3	Q8NFF3 homo sapien
3	85	94.4	211	1 BAK2 HUMAN	Q13014 homo sapien
4	85	94.4	211	1 BAK HUMAN	Q16611 homo sapien
5	85	94.4	211	2 CAG33700	Cag33700 homo sapi
6	84	93.3	163	2 Q8WZS6	Q8WZS6 ovis aries
7	83	92.2	151	2 Q1WXS5	Q1WXS5 mus musculu
8	83	92.2	151	2 AAH57589	AAH57589 mus muscu
9	83	92.2	208	1 BAK MOUSE	O08734 mus musculu
10	83	92.2	209	2 Q8C264	Q8C264 mus musculu
11	83	92.2	209	2 Q9TK59	Q9TK59 rattus norv
12	49	54.4	336	1 K8P3 BACTN	O8A2E9 bacteroides
13	47	52.2	1430	2 Q8PLI2	Q8PLI2 xanthomonas
14	46.5	51.7	357	2 Q8SZ53	Q8SZ53 arabidopsis
15	46	51.1	419	2 Q8AB89	Q8AB89 bacteroides
16	46	51.1	454	2 Q8H7L6	Q8H7L6 phytophthor
17	46	51.1	593	2 Q7VB36	Q7VB36 prochloroco
18	46	51.1	832	1 ATCU SALT1	Q828S4 salmonella
19	46	51.1	832	1 ATCU SALT2	Q82R95 salmonella
20	46	51.1	833	1 ATCU ECO57	Q8X24 escherichia
21	46	51.1	833	1 ATCU ECO11	Q59385 escherichia
22	46	51.1	834	2 Q7C2W2	Q7C2W2 shigella fl
23	46	51.1	834	2 Q83SE2	Q83SE2 shigella fl
24	46	51.1	834	2 Q8FK77	Q8FK77 escherichia
25	46	51.1	915	1 ATCU VIBCH	Q9K275 vibrio chol
26	46	51.1	1226	1 PAT2_CABEL	P34446 caenorhabdi
27	45.5	50.6	260	2 Q89XF1	Q89XF1 bradyrhizob
28	45	50.0	211	2 Q89BY1	Q89BY1 bradyrhizob
29	45	50.0	265	2 Q89QG4	Q89QG4 bradyrhizob
30	45	50.0	297	1 PYRB HELHP	Q7VIT3 helicobacte
31	45	50.0	424	2 Q82SN5	Q82SN5 nitrosomona

32	45	50.0	426	1 ENO_HELPJ	Q9ZMS6 helicobacte
33	45	50.0	426	1 ENO_HELPY	P48285 helicobacte
34	45	50.0	583	2 Q8ZCN7	Q8ZCN7 streptomyce
35	44.5	49.4	532	1 CRTI_SYNY4	P21134 synecocyst
36	44.5	49.4	868	2 Q6CFX2	Q6CFX2 yarrowia li
37	44	48.9	234	2 Q8UJV6	Q8UJV6 agrobacteri
38	44	48.9	258	2 Q9UXV1	Q9UXV1 pyrococcus
39	44	48.9	261	2 O57733	O57733 pyrococcus
40	44	48.9	329	2 Q7D392	Q7D392 agrobacteri
41	44	48.9	329	2 Q8LEF8	Q8LEF8 arabidopsis
42	44	48.9	395	2 Q7ZB37	Q7ZB37 desulfovibr
43	44	48.9	395	2 AAS96278	AAS96278 desulfovibr
44	44	48.9	401	2 Q735C2	Q735C2 bacillus ce
45	44	48.9	401	2 Q81BI1	Q81BI1 bacillus ce

ALIGNMENTS

RESULT 1					
O77738	O77738	PRELIMINARY;	PRT;	80 AA.	
AC	O77738;				
DT	01-NOV-1998 (TREMELrel. 08, Created)				
DT	01-NOV-1998 (TREMELrel. 08, Last sequence update)				
DT	01-JUN-2003 (TREMELrel. 24, Last annotation update)				
DE	Bak protein (Fragment).				
GN	Name=bak;				
OS	Sus scrofa (Pig).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.				
OX	NCBI_TaxID=9823;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Bartling B., Hoffmann J., Holtz J., Schulz R., Heusch G., Darmer D.;				
RL	Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.				
EMBL	A000204; CAA04598.1; -				
DR	GO; GO:0042981; P:regulation of apoptosis; IEA.				
DR	InterPro; IPR000712; BCL2 BH.				
DR	InterPro; IPR002475; BCL2 family.				
DR	Pfam; PF00452; BCL-2; 1.				
DR	SMART; SMO0337; BCL; 1.				
DR	PROSITE; P550062; BCL2_FAMILY; 1.				
DR	PROSITE; PS01259; BH3; 1.				
FT	NON TER 1				
FT	NON TER 80				
SQ	SEQUENCE 80 AA; 8818 MW; FDIAF83BD7D59C86 CRC64;				
Query Match 94.4%; Score 85; DB 2; Length 80;					
Best Local Similarity 100.0%; Pred.No. 9.9e-06;					
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
Qy	1 GQVGRQLAIGDDINRR 17				
Db	23 GQVGRQLAIGDDINRR 39				
RESULT 2					
Q8NFF3	Q8NFF3	PRELIMINARY;	PRT;	190 AA.	
AC	Q8NFF3;				
DT	01-OCT-2002 (TREMELrel. 22, Created)				
DT	01-OCT-2002 (TREMELrel. 22, Last sequence update)				
DT	01-JUN-2003 (TREMELrel. 24, Last annotation update)				
DE	Pro-apoptotic protein BAKM variant.				
GN	Name=BAK;				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Ma J.;				

RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF520590; AM74949.1; -;
DR HSSP: Q16611; 1BXL.
DR GO: GO:0042981; P:regulation of apoptosis; IEA.
DR InterPro: IPR000712; Bcl2_BH.
DR InterPro: IPR002475; Bcl2_family.
DR Pfam: PF00452; Bcl-2; 1.
DR SMART: SM00337; BCL; 1.
DR PROSITE: PS00662; BCL2_FAMILY; 1.
DR PROSITE: PS01080; BH1; 1.
DR PROSITE: PS01258; BH2; 1.
DR PROSITE: PS01259; BH3; 1.
DR PROSITE: PS01259; BH3; 1.
SQ SEQUENCE 190 AA; 21231 MW; A9D4EB8526D0897B CRC64;

Query Match 94.4%; Score 85; DB 2; Length 190;
Best Local Similarity 100.0%; Pred. No. 2.4e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GQVGRQLAIIIGDDINRR 17
DB 51 GQVGRQLAIIIGDDINRR 67

RESULT 3

BAK2_HUMAN STANDARD; PRT; 211 AA.
AC Q13014;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Putative Bcl-2 homologous antagonist/killer 2 (Apoptosis regulator BAK-2).
GN Name=BCL2L7P1; Synonyms=BAK2;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95231654; PubMed=7715731;
RA Kiefer M.C., Brauer M.J., Powers V.C., Wu J.J., Umansky S.R., Tomei L.D., Barr P.J.;
RT "Modulation of apoptosis by the widely distributed Bcl-2 homologue Bak.";
RL Nature 374:736-739(1995).
CC -!- FUNCTION: In the presence of an appropriate stimulus, accelerates programmed cell death by binding to, and antagonizing the a repressor Bcl-2 or its adenovirus homolog E1B 19k protein.
CC -!- SUBUNIT: Forms heterodimers with Bcl-2, E1B 19k protein, and Bcl-2 X(L).
CC -!- SUBCELLULAR LOCATION: Membrane-associated (Potential).
CC -!- TISSUE SPECIFICITY: Expressed in a wide variety of tissues, with highest levels in the heart and skeletal muscle.
CC -!- DOMAIN: Intact BH3 domain is required by BIK, BID, BAK, BAD and BAX for their pro-apoptotic activity and for their interaction with anti-apoptotic members of the Bcl-2 family. Apoptotic members of the Bcl-2 family.
CC -!- SIMILARITY: Belongs to the Bcl-2 family.
CC -!- SIMILARITY: Contains 1 Bcl-2 homology 1 (BH1) domain.
CC -!- SIMILARITY: Contains 1 Bcl-2 homology 2 (BH2) domain.
CC -!- SIMILARITY: Contains 1 Bcl-2 homology 3 (BH3) domain.
CC -!- CAUTION: This is probably the product of a pseudogene.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U16812; AAA74467.1; -;
DR PIR: S58875; S58875.

DR HSSP: Q16611; 1BXL.
DR Genew: HGNC:996; BCL2L7P1.
DR GO: GO:0016020; C:membrane; NAS.
DR GO: GO:0006917; P:induction of apoptosis; NAS.
DR InterPro: IPR000712; Bcl2_BH.
DR InterPro: IPR002475; Bcl2_family.
DR Pfam: PF00452; Bcl-2; 1.
DR SMART: SM00337; BCL; 1.
DR PROSITE: PS00662; BCL2_FAMILY; 1.
DR PROSITE: PS01080; BH1; 1.
DR PROSITE: PS01258; BH2; 1.
DR PROSITE: PS01259; BH3; 1.
DR PROSITE: PS01259; BH3; 1.
KW Apoptosis; Hypothetical protein; Transmembrane.
FT DOMAIN 74 88 BH3.
FT DOMAIN 117 136 BH1.
FT DOMAIN 169 184 BH2.
FT TRANSMEM 188 205 Potential.
SQ SEQUENCE 211 AA; 23411 MW; 703875EC4DCCCLD3 CRC64;

Query Match 94.4%; Score 85; DB 1; Length 211;
Best Local Similarity 100.0%; Pred. No. 2.6e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GQVGRQLAIIIGDDINRR 17
DB 72 GQVGRQLAIIIGDDINRR 88

RESULT 4

BAK_HUMAN STANDARD; PRT; 211 AA.
AC Q16611; Q92533;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Bcl-2 homologous antagonist/killer (Apoptosis regulator BAK) (BCL2-like 7 protein).
GN Name=BAK1; Synonyms=BAK, BCL2L7;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=B-cell;
RX MEDLINE=95231652; PubMed=7715729;
RA Farrow S.N., White J.H.M., Martinou J., Raven T., Pun K.-T., Grinham C.J., Martinou J.C., Brown R.;
RT "Cloning of a bcl-2 homologue by interaction with adenovirus E1B 19K.";
RL Nature 374:731-733(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95231653; PubMed=7715730;
RA Chittenden T., Harrington E.A., O'Connor R., Flemington C., Lutz R.J., Evan G.I., Guild B.C.;
RT "Induction of apoptosis by the Bcl-2 homologue Bak.";
RL Nature 374:733-736(1995).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=95231654; PubMed=7715731;
RA Kiefer M.C., Brauer M.J., Powers V.C., Wu J.J., Umansky S.R., Tomei L.D., Barr P.J.;
RT "Modulation of apoptosis by the widely distributed Bcl-2 homologue Bak.";
RL Nature 374:736-739(1995).
RN [4]
RP SEQUENCE FROM N.A., AND VARIANTS VAL-28 AND ARG-69.
RA Rieder M.J., Livingston R.J., Daniels M.R., Montoya M.A., Chung M.-W., Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D., Schackwitz W.S., Sherwood J.K., Wittrak L.A., Nickerson D.A.;
RT "NIHES-SNPs, environmental genome project, NIHES ES15478, Department of Genome Sciences, Seattle, WA (URL: <http://esg.gs.washington.edu/>).";

Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.

[5]

Sequence from N.A.

MEDLINE=22935763; PubMed=14574404; DOI=10.1039/nature02055;

Mungall A.J., Palmer S.A., Sims S.K., Edwards C.A., Ashurst J.L., Wilming L., Jones M.C., Horton R., Hunt S.E., Scott C.E., Gilbert J.G.R., Clamp M.E., Bethel G., Milne S., Ainscough R., Almeida J.P., Ambrose K.D., Andrews T.D., Ashwell R.I.S., Balade A.K., Baggeley C.L., Bailey J., Banerjee R., Barker D.J., Barlow K.P., Bates K., Beare D.M., Beasley H., Beasley O., Bird C.P., Blakey S.E., Bray-Allen S., Brook J., Brown A.J., Brown J.V., Burford D.C., Burrill W., Burton J., Carder C., Carter N.P., Chapman J.C., Clark S.Y., Clark G., Clee C.M., Clegg S., Cobley V., Collier R.E., Collins J.E., Colman L.K., Corby N.R., Coville G.J., Culley K.M., Dhali P., Davies J., Dunn M., Earthrowl M.B., Ellington A.E., Evans K.A., Faulkner L., Francis M.D., Frankish A., Franklin J., French L., Garner P., Garnett J., Ghori M.J., Gibby L.M., Gillson C.J., Githero R.J., Grafham D.V., Grant M., Gribble S., Griffiths C., Griffiths M.N.D., Hall R., Halls K.S., Hammond S., Harley J.L., Hart E.A., Heath P.D., Heathcote E., Holmes S.J., Howden P.J., Howe K.L., Howell G.R., Huckle E., Humphray S.J., Humphries M.D., Hunt A.R., Johnson C.M., Joy A.A., Kay M., Keenan S.J., Kimberley A.M., King A., Laird G.K., Langford C., Lawlor S., Leongamornlert D.A., Leversha M., Lloyd C.R., Lloyd D.M., Loveland J.E., Lovell J., Martin S., Mashreghi-Mohammadi M., Maslen G.L., Matthews L., McCann O.T., McLaren S.J., McIsaac K., McMurray A., Moore M.J.F., Mullikin J.C., Niblett D., Nickerson T., Novik K.L., Oliver K., Overton-Larty E.K., Parker A., Patel R., Pearce A.V., Peck A.I., Phillimore B.J.C.T., Phillips S., Plumb R.W., Porter K.M., Ramsey A., Rice S.A., Rice C.M., Ross M.T., Seale S.M., Sehra H.K., Sheridan E., Skuce C.D., Smith S., Smith M., Spraggon L., Squares S.L., Steward C.A., Sycamore N., Tamlyn-Hall G., Tester J., Theaker A.J., Thomas D.W., Thorpe A., Tracey A., Tromans A., Tubby B., Wall M., Wallis J.M., West A.P., White S.S., Whitehead S.L., Tubby B., Whittaker H., Wild A., Willey D.J., Wilmer T.S., Wood J.M., Wray P.W., Wyatt J.C., Young L., Younger R.M., Bentley D.R., Coulson A., Durbin R., Hubbard T., Sulston J.E., Dunham J., Rogers J., Beck S.; "The DNA sequence and analysis of human chromosome 6." Nature 425:805-811(2003).

[6]

Sequence from N.A.

Tissue=Lung;

MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Roark S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.M., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A., Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grinchwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.; "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences." Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

[7]

Sequence of 96-206 from N.A.

Eguchi H., Hayashi S.; "Estrogen alters expression of apoptosis-regulators, Bcl-2, Bcl-xL and Bak, as well as susceptibility to therapeutic agents of human breast cancer cells." Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.

[8]

MUTAGENESIS, AND FUNCTION OF BH3 DOMAIN. MEDLINE=96091131; PubMed=8521816;

Chittenden T., Flemington C., Houghton A.B., Ebb R.G., Gallo G.J., Elangovan B., Chinnadurai G., Lutz R.J.; "A conserved domain in Bak, distinct from BH1 and BH2, mediates cell death and protein binding functions." EMBO J. 14:5589-5596(1995).

[9]

STRUCTURE BY NMR OF 72-87. MEDLINE=9712562; PubMed=9020082;

Sattler M., Liang H., Nettelsheim D., Meadows R.P., Harlan J.B., Eberstadt M., Yoon H.S., Shuker S.B., Chang B.S., Minn A.J., Thompson C.B., Fesik S.W.; "Structure of Bcl-xL-Bak peptide complex: recognition between regulators of apoptosis." Science 275:983-986(1997).

!- FUNCTION: In the presence of an appropriate stimulus, accelerates programmed cell death by binding to, and antagonizing the a repressor Bcl-2 or its adenovirus homolog B19 19k protein.

!- SUBUNIT: Forms heterodimers with Bcl-2, ElB 19k protein, and Bcl-x(L).

!- SUBCELLULAR LOCATION: Membrane-bound (Potential).

!- TISSUE SPECIFICITY: Expressed in a wide variety of tissues, with highest levels in the heart and skeletal muscle.

!- DOMAIN: Intact BH3 domain is required by BIK, BID, BAK, BAD and BAX for their pro-apoptotic activity and for their interaction with anti-apoptotic members of the Bcl-2 family.

!- SIMILARITY: Belongs to the Bcl-2 family.

!- SIMILARITY: Contains 1 Bcl-2 homology 1 (BH1) domain.

!- SIMILARITY: Contains 1 Bcl-2 homology 2 (BH2) domain.

!- SIMILARITY: Contains 1 Bcl-2 homology 3 (BH3) domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X84213; CAA58997.1; -

EMBL; U23765; AAA93056.1; -

EMBL; U16811; AAA74466.1; -

EMBL; A1260471; AAO74828.1; -

EMBL; Z93017; CAB65626.1; -

EMBL; BC004431; AAH04431.1; -

EMBL; D88397; BAA13606.1; -

EMBL; D88396; BAA13606.1; JOINED.

PIR; S58873; S58873.

PDB; 1BXL; NMR; B-72-87.

Genew; HGNC:1949; BAK1.

MIM; 600516; -

GO; GO:0008637; P:apoptotic mitochondrial changes; TAS.

InterPro; IPR000712; Bcl2 BH.

InterPro; IPR002475; BCL2_family.

Pfam; PF00452; Bcl-2; 1.

PROSITE; PS00662; BCL2_FAMILY; 1.

PROSITE; PS01080; BH1; 1.

PROSITE; PS01258; BH2; 1.

PROSITE; PS01259; BH3; 1.

3D-structure; Apoptosis; Polymorphism; Transmembrane.

DOMAIN 74 88

DOMAIN 117 136

DOMAIN 169 184

DOMAIN 188 205

TRANSMEM 28 28

VARIANT 69 69

VARIANT 74 75

TURN 76 84

HELIIX 85 85

TURN 211 AA; 23409 MM; A2200FE72A46D04E CRC64;

SEQUENCE 211 AA; 23409 MM; A2200FE72A46D04E CRC64;

Query Match 94.4%; Score 85; DB 1; Length 211;


```

QY 1 GOVGRQLAIIIGDDINR 17
DB 70 GOVGRQLAIIIGDDINR 86

RESULT 8
AAHS7589 PRELIMINARY; PRT; 151 AA.
AC AAHS7589;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Baki protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Strausberg R.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC057589; AAHS7589.1; -.
SQ SEQUENCE 151 AA; 16402 MW; 18C13BFF86E4F33B CRC64;

Query Match 92.2%; Score 83; DB 2; Length 151;
Best Local Similarity 94.1%; Pred. No. 3.9e-05;
Matches 16; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GOVGRQLAIIIGDDINR 17
DB 70 GOVGRQLAIIIGDDINR 86

RESULT 9
BAK MOUSE
ID BAK MOUSE STANDARD; PRT; 208 AA.
AC O08734;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Bcl-2 homologous antagonist/killer (Apoptosis regulator BAK).
DE Name=Bak1; Synonyms=Bak;
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Swiss; TISSUE=Liver;
RX MEDLINE=97446138; PubMed=9299236;
RA Ulrich E., Kaufmann-Zeh A., Hueber A.O., Williamson J.,
RA Chittenden T., Ma A., Ewan G.I.;
RT "Gene structure, cDNA sequence, and expression of murine Bak, a
RT proapoptotic Bcl-2 family member."
RL Genomics 44:195-200(1997).
CC -!- FUNCTION: In the presence of an appropriate stimulus, accelerates
CC programmed cell death by binding to, and antagonizing the a
CC repressor Bcl-2 or its adenovirus homolog E1B 19k protein (By
CC similarity).
CC -!- SUBUNIT: Forms heterodimers with Bcl-2, E1B 19k protein, and Bcl-
CC X(L) (By similarity).
CC -!- SUBCELLULAR LOCATION: Membrane-associated (Potential).
CC -!- TISSUE SPECIFICITY: Widely expressed.
CC -!- DOMAIN: Intact BH3 domain is required by BIK, BID, BAK, BAD and
CC BAX for their pro-apoptotic activity and for their interaction
CC with anti-apoptotic members of the Bcl-2 family (By similarity).
CC -!- SIMILARITY: Belongs to the Bcl-2 family.
CC -!- SIMILARITY: Contains 1 Bcl-2 homology 1 (BH1) domain.
CC -!- SIMILARITY: Contains 1 Bcl-2 homology 2 (BH2) domain.
CC -!- SIMILARITY: Contains 1 Bcl-2 homology 3 (BH3) domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Y13231; CAA73684.1; -.
DR HSP; Q16611; IBL.
DR GMD; MG1:1097161; Bak1.
DR GO; GO:0008635; P: caspase activation via cytochrome c; IMP.
DR InterPro; IPR000712; Bcl2_BH.
DR InterPro; IPR002475; BCL2_family.
DR Pfam; PF00452; Bcl-2; 1.
DR PROSITE; PS00662; BCL2_FAMILY; 1.
DR PROSITE; PS01080; BH1; 1.
DR PROSITE; PS01258; BH2; 1.
DR PROSITE; PS01259; BH3; 1.
KW Apoptosis; Transmembrane.
FT DOMAIN 71 85 BH3.
FT DOMAIN 114 133 BH1.
FT DOMAIN 166 181 BH2.
FT TRANSMEM 185 202 Potential.
SQ SEQUENCE 208 AA; 23300 MW; DAFC11B160C523C9 CRC64;

Query Match 92.2%; Score 83; DB 1; Length 208;
Best Local Similarity 94.1%; Pred. No. 5.4e-05;
Matches 16; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GOVGRQLAIIIGDDINR 17
DB 69 GOVGRQLAIIIGDDINR 85

RESULT 10
Q8C264
ID Q8C264 PRELIMINARY; PRT; 209 AA.
AC Q8C264;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Mus musculus NOD-derived CD11c +ve dendritic cells cDNA, RIKEN full-
DE length enriched library, clone:F630041J23 product:BCL2-
DE antagonist/killer 1, full insert sequence.
GN Name=Bak1;
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

```

```

RN RP SEQUENCE FROM N.A.
RC STRAIN=NOD;
RX MEDLINE=99279253; PubMed=10349636;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
[2]
RN RP SEQUENCE FROM N.A.
RC STRAIN=NOD;
RX MEDLINE=21085660; PubMed=11217851;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
[3]
RN RP SEQUENCE FROM N.A.
RC STRAIN=NOD;
RX MEDLINE=20499374; PubMed=11042159;
RA Carninci P., Shibata Y., Hayatsu M., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
[5]
RN RP SEQUENCE FROM N.A.
RC STRAIN=NOD;
RX MEDLINE=20530913; PubMed=11076861;
RA Shibata K., Itoh M., Aizawa K., Nagaoaka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
[6]
RN RP SEQUENCE FROM N.A.
RC STRAIN=NOD;
RX MEDLINE=20530913; PubMed=11076861;
RA Shibata K., Itoh M., Aizawa K., Nagaoaka S., Sasaki N., Carninci P.,
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hirooka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akashira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (APR-2002) to the ENBL/GenBank/DBJ databases.
DR EMBL: AK089220; BAC40796.1;
DR HSSP: Q16611; 1BXL.
DR GO: GO:0008635; P: caspase activation via cytochrome c; IMP.
DR InterPro: IPR000712; BCL2_BH.
DR InterPro: IPR002475; BCL2_Family.
DR Pfam: PF00452; BCL-2; 1.
DR SMART: SM00337; BCL; 1.
DR PROSITE: PS50062; BCL2_FAMILY; 1.
DR PROSITE: PS01258; BH2; 1.
DR PROSITE: PS01259; BH3; 1.
SQ SEQUENCE 209 AA; 23234 MW; 55E72D29A8AAFD18 CRC64;

```

```

Query Match 92.2%; Score 83; DB 2; Length 209;
Best Local Similarity 94.1%; Pred. No. 5.4e-05;
Matches 16; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GOVGRQLAIGDDINRR 17
Db 70 GOVGRQLAIGDDINRR 86
|||||:|||||
|||||:|||||

RESULT 11
Q9JK59 PRELIMINARY; PRT; 209 AA.
AC Q9JK59;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE BAK protein.
DE Name=Bak;
GN Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=22672518; PubMed=12787069;
RA Itoh T., Itoh A., Pleasure D.;
RT "Bcl-2-related protein family gene expression during oligodendroglial
RT differentiation.";
RL J. Neurochem. 85:1500-1512(2003).
DR EMBL: AF259504; AAF71760.1; -.
DR HSSP: Q16611; 1BXL.
DR GO: GO:0042981; P: regulation of apoptosis; IEA.
DR InterPro: IPR000712; BCL2_BH.
DR InterPro: IPR002475; BCL2_Family.
DR Pfam: PF00452; BCL-2; 1.
DR SMART: SM00337; BCL; 1.
DR PROSITE: PS50062; BCL2_FAMILY; 1.
DR PROSITE: PS01080; BH1; UNKNOWN_1.
DR PROSITE: PS01258; BH2; 1.
DR PROSITE: PS01259; BH3; 1.
SQ SEQUENCE 209 AA; 23153 MW; 2493B814B1972421 CRC64;

Query Match 92.2%; Score 83; DB 2; Length 209;
Best Local Similarity 94.1%; Pred. No. 5.4e-05;
Matches 16; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GOVGRQLAIGDDINRR 17
Db 70 GOVGRQLAIGDDINRR 86
|||||:|||||
|||||:|||||

RESULT 12
K6P3 BACTN STANDARD; PRT; 336 AA.
ID K6P3 BACTN
AC Q8A2E9;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE 6-phosphofructokinase 3 (EC 2.7.1.11) (Phosphofructokinase 3)
DE (Phosphohexokinase 3).
GN Name=PfkF3; OrderedLocNames=ET3356;
OS Bacteroides thetaiotaomicron.
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Bacteroidaceae; Bacteroides.
OX NCBI_TaxID=818;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VPI-5482 / ATCC 29148;
RX MEDLINE=22550858; PubMed=12663928; DOI=10.1126/science.1080029;
RA Xu J., Bjursell M.K., Himrod J., Deng S., Carmichael L.K.,
RA Chiang H.C., Hooper L.V., Gordon J.I.;
RT "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis.";

```

RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities";
RL Nature 417:459-463 (2002).
DR EMBL: AEC01181; AM36678.1; -.
DR InterPro: IPR008619; F11_haemagg.
DR Pfam: PF05594; F11_haemagg; 5.
KW Complete proteome.
SQ SEQUENCE 1430 AA; 145025 MW; 6BA2BCD1BB27E053 CRC64;

Query Watch 52.2%; Score 47; DB 2; Length 1430;
Best Local Similarity 52.9%; Pred. NO. 2.2e+02;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps

QY 1 GOVGRQLAIIGDINRR 17
|||||:|:|:|:
DB 1218 GOVGATTALAGDQFNQ 1234

RESULT 14
Q9SZ53
ID Q9SZ53 PRELIMINARY; PRT; 357 AA.
AC Q9SZ53;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DT Protein phosphatase 2C-like protein (AT4G31860/F11C18_60).
GN Name=AT4G31860;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; euclityledons; Core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Terry N., Ardiles W., Buysshaert C., Dasseville R.,
RA De Clerck R., De Keyser A., Neyt P., Rouze P., Van Den Daele H.,
RA Villaroel R., Gielen J., Van Montagu M., Hobeisel J., Mewes H.W.,
RA Lemcke K., Mayer K.F.X.,
RA Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RL [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.K., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RA Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RL [4]
RP SEQUENCE FROM N.A.
RA Kim C.J., Chen H., Cheuk R., Shinn P., Banh J., Bowser L.,
RA Carninci P., Chang E., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,
RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Wu H.C., Yamada K., Yamamura Y., Yu G., Yu S.,
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RA Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Terry N., Ardiles W., Buysshaert C., Dasseville R., De Clerck R.,
RA De Keyser A., Neyt P., Rouze P., Van Den Daele H., Villaroel R.,
RA Gielen J., Van Montagu M., Mewes H.W., Lemcke K., Mayer K.F.X.;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL049607; CAB40756.1; -.
RN [6]

Search completed: November 10, 2004, 13:38:48
Job time : 23.9793 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 12:32:37 ; Search time 8.981 Seconds
(without alignments)
191.991 Million cell updates/sec

Title: US-10-092-750-42

Perfect score: 123

Sequence: 1 GVSEAGTFTPLFLGIASRLRSA 26

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	38.2	505	4	US-09-877-476-10
2	47	38.2	505	4	US-09-877-476-18
3	47	38.2	506	3	US-08-888-998-2
4	47	38.2	506	3	US-09-362-633-2
5	47	38.2	506	4	US-09-877-476-2
6	47	38.2	506	4	US-09-877-476-8
7	47	38.2	506	4	US-09-877-476-12
8	47	38.2	506	4	US-09-877-476-14
9	47	38.2	506	4	US-09-877-476-16
10	47	38.2	506	4	US-09-877-476-20
11	47	38.2	506	4	US-09-877-476-22
12	47	38.2	506	4	US-09-877-476-24
13	47	38.2	506	4	US-09-877-476-26
14	47	38.2	506	4	US-09-877-476-30
15	47	38.2	506	4	US-09-877-476-36
16	47	38.2	506	4	US-09-877-476-38
17	47	38.2	506	4	US-09-877-476-40
18	47	38.2	506	4	US-08-309-512-10
19	47	38.2	740	5	PCT-US92-08756A-10
20	45	36.6	211	4	US-09-198-452A-902
21	45	36.6	765	1	US-08-309-512-5
22	45	36.6	765	5	PCT-US92-08756A-5
23	44	35.8	885	4	US-09-248-996A-14427
24	43	35.0	74	4	US-09-252-991A-22657
25	43	35.0	78	4	US-09-621-976-5620
26	43	35.0	85	4	US-09-270-767-46286
27	43	35.0	113	4	US-09-886-319A-25

28 43 35.0 113 4 US-09-886-319A-26 Sequence 26, Appl
29 43 35.0 113 4 US-09-513-998C-7824 Sequence 7824, Ap
30 43 35.0 410 4 US-09-489-039A-7598 Sequence 7598, Ap
31 43 35.0 740 4 US-09-252-991A-20586 Sequence 20586, A
32 42 34.1 162 4 US-09-252-991A-23097 Sequence 23097, A
33 42 34.1 167 4 US-09-489-039A-9260 Sequence 9260, Ap
34 42 34.1 372 1 US-08-597-236-8 Sequence 8, Appl
35 42 34.1 372 1 US-08-746-682A-8 Sequence 8, Appl
36 42 34.1 394 4 US-09-248-796A-22220 Sequence 22220, A
37 42 34.1 493 3 US-09-117-250-3 Sequence 3, Appl
38 41.5 33.7 190 3 US-08-078-271B-1 Sequence 1, Appl
39 41.5 33.7 374 1 US-08-464-148-2 Sequence 2, Appl
40 41.5 33.7 374 1 US-08-385-500-2 Sequence 2, Appl
41 41.5 33.7 374 1 US-08-846-784-2 Sequence 2, Appl
42 41 33.3 59 4 US-09-621-976-5289 Sequence 5289, Ap
43 41 33.3 80 4 US-09-489-039A-11340 Sequence 11340, A
44 41 33.3 91 4 US-09-621-976-6683 Sequence 6683, Ap
45 41 33.3 115 4 US-09-325-932A-53 Sequence 53, Appl

ALIGNMENTS

RESULT 1
US-09-877-476-10
; Sequence 10, Application US/09877476
; Patent No. 6713664
; GENERAL INFORMATION:
; APPLICANT: Jaworski, Jan G.
; APPLICANT: Blacklock, Brenda J.
; TITLE OF INVENTION: FATTY ACID ELONGASE 3-KETOACYL COA
; FILE REFERENCE: 07148-108001
; CURRENT APPLICATION NUMBER: US/09/877,476
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 60/210,326
; PRIOR FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: 5' 74 amino acids from A. thaliana PAEL (SEQ ID
; OTHER INFORMATION: NO:2) and 3' 431 amino acids from B. napus
; OTHER INFORMATION: elongase KCS (SEQ ID NO:4); designated At74
US-09-877-476-10
Query Match 38.2%; Score 47; DB 4; Length 505;
Best Local Similarity 71.4%; Pred. No. 19;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 9 FFLSTFLGIASRL 22
|||:|||||
Db 23 FLTAFLAGKASRL 36
RESULT 2
US-09-877-476-18
; Sequence 18, Application US/09877476
; Patent No. 6713664
; GENERAL INFORMATION:
; APPLICANT: Jaworski, Jan G.
; APPLICANT: Blacklock, Brenda J.
; TITLE OF INVENTION: FATTY ACID ELONGASE 3-KETOACYL COA
; FILE REFERENCE: 07148-108001
; CURRENT APPLICATION NUMBER: US/09/877,476
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 60/210,326
; PRIOR FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 56

SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 18
LENGTH: 505
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: 5' 74 amino acids from A. thaliana PAEI (SEQ ID NO:2) and 3' 431 amino acids from B. napus
OTHER INFORMATION: eLongase KCS (SEQ ID NO:4) having a mutation at residue 306; designated At74 G306D; hypothetical
US-09-877-476-18

Query Match 38.2%; Score 47; DB 4; Length 505;
Best Local Similarity 71.4%; Pred. No. 19;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 9 FPLSTFLGLIASRL 22
|||: ||| ||||
Db 23 FPLTAFLAGKASRL 36

RESULT 3
US-08-888-998-2
Sequence 2, Application US/08888998
Patent No. 6124524
GENERAL INFORMATION:
APPLICANT: JAMES, Douglas W.
APPLICANT: LIM, Eda
APPLICANT: KELLER, Janis
APPLICANT: DOONER, Hugo K.
TITLE OF INVENTION: FAEL GENES AND THEIR USES
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Khourie and Crew
STREET: Steuart Street Tower, One Market Plaza
CITY: San Francisco
STATE: California
COUNTRY: US
ZIP: 94105-1493
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/888,998
FILING DATE: 07-JUL-1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/329,603
FILING DATE: 26-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Bastian, Kevin L.
REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 12176-004300
TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 506 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Protein
LOCATION: 1..506
OTHER INFORMATION: /note= "Amino acid sequence of FAEL
OTHER INFORMATION: protein."
US-08-888-998-2

Query Match 38.2%; Score 47; DB 3; Length 506;

Best Local Similarity 71.4%; Pred. No. 19;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 9 FPLSTFLGLIASRL 22
|||: ||| ||||
Db 23 FPLTAFLAGKASRL 36

RESULT 4
US-09-362-633-2
Sequence 2, Application US/09362633
Patent No. 6184355
GENERAL INFORMATION:
APPLICANT: JAMES, Douglas W.
APPLICANT: LIM, Eda
APPLICANT: KELLER, Janis
APPLICANT: DOONER, Hugo K.
TITLE OF INVENTION: FAEL GENES AND THEIR USES
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Khourie and Crew
STREET: Steuart Street Tower, One Market Plaza
CITY: San Francisco
STATE: California
COUNTRY: US
ZIP: 94105-1493
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/362,633
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/888,998
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Bastian, Kevin L.
REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 12176-004300
TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 506 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Protein
LOCATION: 1..506
OTHER INFORMATION: /note= "Amino acid sequence of FAEL
OTHER INFORMATION: protein."
US-09-362-633-2

Query Match 38.2%; Score 47; DB 3; Length 506;
Best Local Similarity 71.4%; Pred. No. 19;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 9 FPLSTFLGLIASRL 22
|||: ||| ||||
Db 23 FPLTAFLAGKASRL 36

RESULT 5
US-09-877-476-2
Sequence 2, Application US/09877476
Patent No. 6713664
GENERAL INFORMATION:


```

; GENERAL INFORMATION:
; APPLICANT: Jaworski, Jan G.
; APPLICANT: Blacklock, Brenda J.
; TITLE OF INVENTION: FATTY ACID ELONGASE 3-KETOACYL COA
; TITLE OF INVENTION: SYNTHASE POLYPEPTIDES
; FILE REFERENCE: 07148-108001

```

; CURRENT APPLICATION NUMBER: US/09/877,476
 ; CURRENT FILING DATE: 2001-06-08
 ; PRIOR APPLICATION NUMBER: US 60/210,326
 ; PRIOR FILING DATE: 2000-06-08
 ; NUMBER OF SEQ ID NOS: 56
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 16
 ; LENGTH: 506
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: 5' 114 amino acids from A. thaliana PAE1 (SEQ ID
 ; OTHER INFORMATION: NO:2) and 3' 392 amino acids from B. napus
 ; OTHER INFORMATION: elongase KCS (SEQ ID NO:4) having mutation at
 ; OTHER INFORMATION: residue 307; designated At114 G307D; hypothetical
 ; OTHER INFORMATION: residue 307; designated At114 G307D; hypothetical
 US-09-877-476-16

Query Match 38.2%; Score 47; DB 4; Length 506;
 Best Local Similarity 71.4%; Pred. No. 19;
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 9 FPLSTFLGIASRL 22
 |||: || |||||
 Db 23 FPLTAFLAGKASRL 36

RESULT 10

US-09-877-476-20
 ; Sequence 20, Application US/09877476
 ; Patent No. 6713664
 ; GENERAL INFORMATION:
 ; APPLICANT: Jaworski, Jan G.
 ; APPLICANT: Blacklock, Brenda J.
 ; TITLE OF INVENTION: FATTY ACID ELONGASE 3-KETOACYL COA
 ; TITLE OF INVENTION: SYNTHASE POLYPEPTIDES
 ; FILE REFERENCE: 07148-108001

; CURRENT APPLICATION NUMBER: US/09/877,476
 ; CURRENT FILING DATE: 2001-06-08
 ; PRIOR APPLICATION NUMBER: US 60/210,326
 ; PRIOR FILING DATE: 2000-06-08
 ; NUMBER OF SEQ ID NOS: 56
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 20
 ; LENGTH: 506
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: 5' 114 amino acids from A. thaliana PAE1 (SEQ ID
 ; OTHER INFORMATION: NO:2) and 3' 392 amino acids from B. napus
 ; OTHER INFORMATION: elongase KCS (SEQ ID NO:4) having mutations at
 ; OTHER INFORMATION: positions 91, 92 and 307; designated At114 L91C
 ; OTHER INFORMATION: K92R G307D; hypothetical

US-09-877-476-20

Query Match 38.2%; Score 47; DB 4; Length 506;
 Best Local Similarity 71.4%; Pred. No. 19;
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 9 FPLSTFLGIASRL 22
 |||: || |||||
 Db 23 FPLTAFLAGKASRL 36

RESULT 11

US-09-877-476-22
 ; Sequence 22, Application US/09877476
 ; Patent No. 6713664
 ; GENERAL INFORMATION:
 ; APPLICANT: Jaworski, Jan G.
 ; APPLICANT: Blacklock, Brenda J.
 ; TITLE OF INVENTION: FATTY ACID ELONGASE 3-KETOACYL COA
 ; TITLE OF INVENTION: SYNTHASE POLYPEPTIDES
 ; FILE REFERENCE: 07148-108001

; CURRENT APPLICATION NUMBER: US/09/877,476
 ; CURRENT FILING DATE: 2001-06-08
 ; PRIOR APPLICATION NUMBER: US 60/210,326
 ; PRIOR FILING DATE: 2000-06-08
 ; NUMBER OF SEQ ID NOS: 56
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 22
 ; LENGTH: 506
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: 5' 114 amino acids from A. thaliana PAE1 (SEQ ID
 ; OTHER INFORMATION: NO:2) and 3' 392 amino acids from B. napus
 ; OTHER INFORMATION: elongase KCS (SEQ ID NO:4) having mutations at
 ; OTHER INFORMATION: positions 92 and 307; designated At114 K92R G307D;
 ; OTHER INFORMATION: hypothetical
 US-09-877-476-22

Query Match 38.2%; Score 47; DB 4; Length 506;
 Best Local Similarity 71.4%; Pred. No. 19;
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 9 FPLSTFLGIASRL 22
 |||: || |||||
 Db 23 FPLTAFLAGKASRL 36

RESULT 12

US-09-877-476-24
 ; Sequence 24, Application US/09877476
 ; Patent No. 6713664
 ; GENERAL INFORMATION:
 ; APPLICANT: Jaworski, Jan G.
 ; APPLICANT: Blacklock, Brenda J.

; TITLE OF INVENTION: FATTY ACID ELONGASE 3-KETOACYL COA
 ; TITLE OF INVENTION: SYNTHASE POLYPEPTIDES
 ; FILE REFERENCE: 07148-108001
 ; CURRENT APPLICATION NUMBER: US/09/877,476
 ; CURRENT FILING DATE: 2001-06-08

; PRIOR APPLICATION NUMBER: US 60/210,326
 ; PRIOR FILING DATE: 2000-06-08
 ; NUMBER OF SEQ ID NOS: 56
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 24
 ; LENGTH: 506
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: 5' 254 amino acids from A. thaliana PAE1 (SEQ ID
 ; OTHER INFORMATION: NO:2) and 3' 252 amino acids from B. napus
 ; OTHER INFORMATION: elongase KCS (SEQ ID NO:4); designated At254

US-09-877-476-24

Query Match 38.2%; Score 47; DB 4; Length 506;
 Best Local Similarity 71.4%; Pred. No. 19;
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 9 FPLSTFLGIASRL 22
 |||: || |||||
 Db 23 FPLTAFLAGKASRL 36

RESULT 13

US-09-877-476-26
 ; Sequence 26, Application US/09877476
 ; Patent No. 6713664
 ; GENERAL INFORMATION:
 ; APPLICANT: Jaworski, Jan G.

; APPLICANT: Blacklock, Brenda J.
 ; TITLE OF INVENTION: FATTY ACID ELONGASE 3-KETOACYL COA
 ; TITLE OF INVENTION: SYNTHASE POLYPEPTIDES
 ; FILE REFERENCE: 07148-108001
 ; CURRENT APPLICATION NUMBER: US/09/877,476

```
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 60/210,326
; PRIOR FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 506
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: 5' 173 amino acids from A. thaliana FAE1 (SEQ ID
; OTHER INFORMATION: NO:2) and 3' 333 amino acids from B. napus
; OTHER INFORMATION: elongase KCS (SEQ ID NO:4); designated At173
; NAME/KEY: VARIANT
; LOCATION: (0)...(0)
; OTHER INFORMATION: Xaa = Pro or Gln
US-09-877-476-26

Query Match      38.2%; Score 47; DB 4; Length 506;
Best Local Similarity 71.4%; Pred. No. 19;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      9 FPLSTFLGLIASRL 22
      |||: ||| |||
Db      23 FPLTAFLAGKASRL 36

RESULT 14
US-09-877-476-30
; Sequence 30, Application US/09877476
; Patent No. 6713664
; GENERAL INFORMATION:
; APPLICANT: Jaworski, Jan G.
; APPLICANT: Blacklock, Brenda J.
; TITLE OF INVENTION: FATTY ACID ELONGASE 3-KETOACYL COA
; FILE REFERENCE: 07148-108001
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US/09/877,476
; PRIOR FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 506
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: 5' 399 amino acids from A. thaliana FAE1 (SEQ ID
; OTHER INFORMATION: NO:2) and 3' 107 amino acids from B. napus
; OTHER INFORMATION: elongase KCS (SEQ ID NO:4); designated At399
; NAME/KEY: VARIANT
; LOCATION: (0)...(0)
; OTHER INFORMATION: Xaa = Pro or Gln
US-09-877-476-30

Query Match      38.2%; Score 47; DB 4; Length 506;
Best Local Similarity 71.4%; Pred. No. 19;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      9 FPLSTFLGLIASRL 22
      |||: ||| |||
Db      23 FPLTAFLAGKASRL 36

RESULT 15
US-09-877-476-36
; Sequence 36, Application US/09877476
; Patent No. 6713664
; GENERAL INFORMATION:
; APPLICANT: Jaworski, Jan G.
; APPLICANT: Blacklock, Brenda J.
; TITLE OF INVENTION: FATTY ACID ELONGASE 3-KETOACYL COA
```

```
; TITLE OF INVENTION: SYNTHASE POLYPEPTIDES
; FILE REFERENCE: 07148-108001
; CURRENT APPLICATION NUMBER: US/09/877,476
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 60/210,326
; PRIOR FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36
; LENGTH: 506
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: 506 amino acids from A. thaliana FAE1 (SEQ ID
; OTHER INFORMATION: NO:2) having a mutation at residue 92; designated
; OTHER INFORMATION: At K92R; hypothetical
US-09-877-476-36

Query Match      38.2%; Score 47; DB 4; Length 506;
Best Local Similarity 71.4%; Pred. No. 19;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      9 FPLSTFLGLIASRL 22
      |||: ||| |||
Db      23 FPLTAFLAGKASRL 36
```

Search completed: November 10, 2004, 13:44:11
Job time : 8.981 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 16:36:12 ; Search time 27.9309 Seconds
(without alignments)
328.807 Million cell updates/sec

Title: US-10-092-750-42

Perfect score: 123

Sequence: 1 GVSEAGTPTLSTFLGIASRLRSVA 26

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

- Database : Published Applications AA:*
- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
 - 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
 - 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
 - 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
 - 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
 - 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	123	100.0	26	14	US-10-092-750-42
2	51	41.5	87	17	US-10-425-115-264584
3	50	40.7	264	15	US-10-424-599-155062
4	48	39.0	894	16	US-10-437-963-119916
5	47	38.2	505	9	US-09-877-476-10
6	47	38.2	505	9	US-09-877-476-18
7	47	38.2	505	16	US-10-758-524-10
8	47	38.2	505	16	US-10-758-524-18
9	47	38.2	506	9	US-09-877-476-2
10	47	38.2	506	9	US-09-877-476-8
11	47	38.2	506	9	US-09-877-476-12
12	47	38.2	506	9	US-09-877-476-14
13	47	38.2	506	9	US-09-877-476-16

14	47	38.2	506	9	US-09-877-476-20	Sequence 20, Appl
15	47	38.2	506	9	US-09-877-476-22	Sequence 22, Appl
16	47	38.2	506	9	US-09-877-476-24	Sequence 24, Appl
17	47	38.2	506	9	US-09-877-476-26	Sequence 26, Appl
18	47	38.2	506	9	US-09-877-476-30	Sequence 30, Appl
19	47	38.2	506	9	US-09-877-476-36	Sequence 36, Appl
20	47	38.2	506	9	US-09-877-476-38	Sequence 38, Appl
21	47	38.2	506	9	US-09-877-476-40	Sequence 40, Appl
22	47	38.2	506	15	US-10-276-977-7	Sequence 7, Appl
23	47	38.2	506	16	US-10-758-524-2	Sequence 2, Appl
24	47	38.2	506	16	US-10-758-524-8	Sequence 8, Appl
25	47	38.2	506	16	US-10-758-524-12	Sequence 12, Appl
26	47	38.2	506	16	US-10-758-524-14	Sequence 14, Appl
27	47	38.2	506	16	US-10-758-524-16	Sequence 16, Appl
28	47	38.2	506	16	US-10-758-524-20	Sequence 20, Appl
29	47	38.2	506	16	US-10-758-524-22	Sequence 22, Appl
30	47	38.2	506	16	US-10-758-524-24	Sequence 24, Appl
31	47	38.2	506	16	US-10-758-524-26	Sequence 26, Appl
32	47	38.2	506	16	US-10-758-524-30	Sequence 30, Appl
33	47	38.2	506	16	US-10-758-524-36	Sequence 36, Appl
34	47	38.2	506	16	US-10-758-524-38	Sequence 38, Appl
35	47	38.2	506	16	US-10-758-524-40	Sequence 40, Appl
36	47	38.2	819	16	US-10-437-963-151529	Sequence 151529,
37	47	38.2	1042	16	US-10-437-963-151525	Sequence 151525,
38	46.5	37.8	231	15	US-10-425-114-65092	Sequence 65092, A
39	46.5	37.8	259	15	US-10-424-599-146049	Sequence 146049,
40	46	37.4	217	17	US-10-425-115-355887	Sequence 355887,
41	46	37.4	376	15	US-10-282-122A-76580	Sequence 76580, A
42	46	37.4	485	9	US-09-741-669-299	Sequence 299, App
43	46	37.4	605	14	US-10-156-761-9470	Sequence 9470, Ap
44	45	36.6	96	17	US-10-425-115-217978	Sequence 217978,
45	45	36.6	152	17	US-10-425-115-240412	Sequence 240412,

ALIGNMENTS

RESULT 1
US-10-092-750-42
; Sequence 42, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; APPLICANT: Wright, Martin C.
; TITLE OF INVENTION: Polypeptides Interactive with BCL-XI
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092,750
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/274,526
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-42

Query Match 100.0%; Score 123; DB 14; Length 26;
Best Local Similarity 100.0%; Pred. No. 2.4e-12;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVSEAGTPTLSTFLGIASRLRSVA 26
Db 1 GVSEAGTPTLSTFLGIASRLRSVA 26

RESULT 2
US-10-425-115-264584
; Sequence 264584, Application US/10425115
; Publication No. US2004021472A1
; GENERAL INFORMATION:

```

; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 264584
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_172912C.1.pep
US-10-425-115-264584

```

```

Query Match          41.5%; Score 51; DB 17; Length 87;
Best Local Similarity 44.8%; Pred. No. 3;
Matches 13; Conservative 2; Mismatches 4; Indels 10; Gaps 1;

```

```

QY 5 AEGTFPLSTF-----LLGIASRLR 23
|||:|||||
DB 50 AEGVFLRTFGGDEPQQVLLGVQRMR 78
|||:|||||

```

```

RESULT 3
US-10-424-599-155062
; Sequence 155062, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 155062
; LENGTH: 264
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_111042C.1.pep
US-10-424-599-155062

```

```

Query Match          40.7%; Score 50; DB 15; Length 264;
Best Local Similarity 62.5%; Pred. No. 16;
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 8 TPLSTFLGIASRLR 23
|||:|||||
DB 21 TPLTLLLLGIVSRIR 36
|||:|||||

```

```

RESULT 4
US-10-437-963-119916
; Sequence 119916, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

```

```

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 119916
; LENGTH: 894
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_23086C.1.pep
US-10-437-963-119916

```

```

Query Match          39.0%; Score 48; DB 16; Length 894;
Best Local Similarity 45.0%; Pred. No. 1.4e+02;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

```

```

QY 3 SEAGTFPLSTELLGIASRL 22
|||:|||||
DB 88 ANAPARLPSEFWLGLANKL 107
|||:|||||

```

```

RESULT 5
US-09-877-476-10
; Sequence 10, Application US/09877476
; Patent No. US20020049994A1
; GENERAL INFORMATION:
; APPLICANT: Jaworski, Jan G.
; APPLICANT: Blacklock, Brenda J.
; TITLE OF INVENTION: FATTY ACID ELONGASE 3-KETOACYL COA
; TITLE OF INVENTION: SYNTHASE POLYPEPTIDES
; FILE REFERENCE: 07148-108001
; CURRENT APPLICATION NUMBER: US/09/877,476
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 60/210,326
; PRIOR FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: 5' 74 amino acids from A. thaliana PAE1 (SEQ ID
; OTHER INFORMATION: NO:2) and 3' 431 amino acids from B. napus
; OTHER INFORMATION: elongase KCS (SEQ ID NO:4); designated At74
US-09-877-476-10

```

```

Query Match          38.2%; Score 47; DB 9; Length 505;
Best Local Similarity 71.4%; Pred. No. 1e+02;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 9 FPLSTFLGIASRL 22
|||:|||||
DB 23 FPLTAFLAGASRL 36
|||:|||||

```

```

RESULT 6
US-09-877-476-18
; Sequence 18, Application US/09877476
; Patent No. US20020049994A1
; GENERAL INFORMATION:
; APPLICANT: Jaworski, Jan G.
; APPLICANT: Blacklock, Brenda J.
; TITLE OF INVENTION: FATTY ACID ELONGASE 3-KETOACYL COA
; TITLE OF INVENTION: SYNTHASE POLYPEPTIDES
; FILE REFERENCE: 07148-108001
; CURRENT APPLICATION NUMBER: US/09/877,476
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 60/210,326
; PRIOR FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0

```

```
; SEQ ID NO 18
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: 5' 74 amino acids from A. thaliana FAE1 (SEQ ID
; OTHER INFORMATION: NO:2) and 3' 431 amino acids from B. napus
; OTHER INFORMATION: elongase KCS (SEQ ID NO:4) having a mutation at
; OTHER INFORMATION: residue 306; designated At74 G306D; hypothetical
US-09-877-476-18

Query Match      38.2%; Score 47; DB 9; Length 505;
Best Local Similarity 71.4%; Pred. No. 1e+02;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy      9 FPLSTFLLGIASRL 22
Db      23 FPLTAFLAGKASRL 36

RESULT 7
US-10-758-524-10
; Sequence 10, Application US/10758524
; Publication No. US20040139498A1
; GENERAL INFORMATION:
; APPLICANT: Jaworski, Jan G.
; TITLE OF INVENTION: SATTY ACID ELONGASE 3-KETOACYL COA
; FILE REFERENCE: 07148-108002
; CURRENT APPLICATION NUMBER: US/10/758,524
; CURRENT FILING DATE: 2004-01-15
; PRIOR FILING DATE: 09/877,476
; PRIOR FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: 60/210,326
; PRIOR FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: 5' 74 amino acids from A. thaliana FAE1 (SEQ ID
; OTHER INFORMATION: NO:2) and 3' 431 amino acids from B. napus
; OTHER INFORMATION: elongase KCS (SEQ ID NO:4); designated At74
US-10-758-524-10

Query Match      38.2%; Score 47; DB 16; Length 505;
Best Local Similarity 71.4%; Pred. No. 1e+02;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy      9 FPLSTFLLGIASRL 22
Db      23 FPLTAFLAGKASRL 36

RESULT 8
US-10-758-524-18
; Sequence 18, Application US/10758524
; Publication No. US20040139498A1
; GENERAL INFORMATION:
; APPLICANT: Jaworski, Jan G.
; TITLE OF INVENTION: SATTY ACID ELONGASE 3-KETOACYL COA
; FILE REFERENCE: 07148-108002
; CURRENT APPLICATION NUMBER: US/10/758,524
; CURRENT FILING DATE: 2004-01-15
; PRIOR FILING DATE: 09/877,476
; PRIOR APPLICATION NUMBER: 60/210,326
; PRIOR FILING DATE: 2000-06-08
```

```
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: 5' 74 amino acids from A. thaliana FAE1 (SEQ ID
; OTHER INFORMATION: NO:2) and 3' 431 amino acids from B. napus
; OTHER INFORMATION: elongase KCS (SEQ ID NO:4) having a mutation at
; OTHER INFORMATION: residue 306; designated At74 G306D; hypothetical
US-10-758-524-18

Query Match      38.2%; Score 47; DB 16; Length 505;
Best Local Similarity 71.4%; Pred. No. 1e+02;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy      9 FPLSTFLLGIASRL 22
Db      23 FPLTAFLAGKASRL 36

RESULT 9
US-09-877-476-2
; Sequence 2, Application US/09877476
; Patent No. US20020049994A1
; GENERAL INFORMATION:
; APPLICANT: Jaworski, Jan G.
; TITLE OF INVENTION: SATTY ACID ELONGASE 3-KETOACYL COA
; FILE REFERENCE: 07148-108001
; CURRENT APPLICATION NUMBER: US/09/877,476
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 60/210,326
; PRIOR FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 506
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-877-476-2

Query Match      38.2%; Score 47; DB 9; Length 506;
Best Local Similarity 71.4%; Pred. No. 1e+02;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy      9 FPLSTFLLGIASRL 22
Db      23 FPLTAFLAGKASRL 36

RESULT 10
US-09-877-476-8
; Sequence 8, Application US/09877476
; Patent No. US20020049994A1
; GENERAL INFORMATION:
; APPLICANT: Jaworski, Jan G.
; TITLE OF INVENTION: SATTY ACID ELONGASE 3-KETOACYL COA
; FILE REFERENCE: 07148-108001
; CURRENT APPLICATION NUMBER: US/09/877,476
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 60/210,326
; PRIOR FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 506
; TYPE: PRT
; ORGANISM: Artificial Sequence
```

```

; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (0)...(0)
; OTHER INFORMATION: Xaa = Ala or Thr
; OTHER INFORMATION: 5' 114 amino acids from A. thaliana FAE1 (SEQ ID
; OTHER INFORMATION: NO:2) and 3' 392 amino acids from B. napus
; OTHER INFORMATION: elongase KCS (SEQ ID NO:4); designated At114
; OTHER INFORMATION: elongase KCS (SEQ ID NO:4); designated At114 K92R
US-09-877-476-8

```

```

Query Match      38.2%; Score 47; DB 9; Length 506;
Best Local Similarity 71.4%; Pred. No. 1e+02;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 9 FPLSTFLGIASRL 22
   |||: || |||||
DB 23 FPLTAFLAGKASRL 36

```

```

RESULT 11
US-09-877-476-12
; Sequence 12, Application US/09877476
; Patent No. US20020049994A1
; GENERAL INFORMATION:
; APPLICANT: Jaworski, Jan G.
; APPLICANT: Blacklock, Brenda J.
; TITLE OF INVENTION: FATTY ACID ELONGASE 3-KETOACYL COA
; TITLE OF INVENTION: SYNTHASE POLYPEPTIDES
; FILE REFERENCE: 07148-108001
; CURRENT APPLICATION NUMBER: US/09/877,476
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 60/210,326
; PRIOR FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 506
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: 5' 114 amino acids from A. thaliana FAE1 (SEQ ID
; OTHER INFORMATION: NO:2) and 3' 392 amino acids from B. napus
; OTHER INFORMATION: elongase KCS (SEQ ID NO:4) having mutations at
; OTHER INFORMATION: residues 91 and 92; designated At114 L91C K92R
US-09-877-476-12

```

```

Query Match      38.2%; Score 47; DB 9; Length 506;
Best Local Similarity 71.4%; Pred. No. 1e+02;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 9 FPLSTFLGIASRL 22
   |||: || |||||
DB 23 FPLTAFLAGKASRL 36

```

```

RESULT 12
US-09-877-476-14
; Sequence 14, Application US/09877476
; Patent No. US20020049994A1
; GENERAL INFORMATION:
; APPLICANT: Jaworski, Jan G.
; APPLICANT: Blacklock, Brenda J.
; TITLE OF INVENTION: FATTY ACID ELONGASE 3-KETOACYL COA
; TITLE OF INVENTION: SYNTHASE POLYPEPTIDES
; FILE REFERENCE: 07148-108001
; CURRENT APPLICATION NUMBER: US/09/877,476
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 60/210,326
; PRIOR FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 506
; TYPE: PRT
; ORGANISM: Artificial Sequence

```

```

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: 5' 114 amino acids from A. thaliana FAE1 (SEQ ID
; OTHER INFORMATION: NO:2) and 3' 392 amino acids from B. napus
; OTHER INFORMATION: elongase KCS (SEQ ID NO:4); having a mutation at
; OTHER INFORMATION: position 92; designated At114 K92R
US-09-877-476-14

```

```

Query Match      38.2%; Score 47; DB 9; Length 506;
Best Local Similarity 71.4%; Pred. No. 1e+02;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 9 FPLSTFLGIASRL 22
   |||: || |||||
DB 23 FPLTAFLAGKASRL 36

```

```

RESULT 13
US-09-877-476-16
; Sequence 16, Application US/09877476
; Patent No. US20020049994A1
; GENERAL INFORMATION:
; APPLICANT: Jaworski, Jan G.
; APPLICANT: Blacklock, Brenda J.
; TITLE OF INVENTION: FATTY ACID ELONGASE 3-KETOACYL COA
; TITLE OF INVENTION: SYNTHASE POLYPEPTIDES
; FILE REFERENCE: 07148-108001
; CURRENT APPLICATION NUMBER: US/09/877,476
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 60/210,326
; PRIOR FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 506
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: 5' 114 amino acids from A. thaliana FAE1 (SEQ ID
; OTHER INFORMATION: NO:2) and 3' 392 amino acids from B. napus
; OTHER INFORMATION: elongase KCS (SEQ ID NO:4) having mutation at
; OTHER INFORMATION: residue 307; designated At114 G307D; hypothetical
US-09-877-476-16

```

```

Query Match      38.2%; Score 47; DB 9; Length 506;
Best Local Similarity 71.4%; Pred. No. 1e+02;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 9 FPLSTFLGIASRL 22
   |||: || |||||
DB 23 FPLTAFLAGKASRL 36

```

```

RESULT 14
US-09-877-476-20
; Sequence 20, Application US/09877476
; Patent No. US20020049994A1
; GENERAL INFORMATION:
; APPLICANT: Jaworski, Jan G.
; APPLICANT: Blacklock, Brenda J.
; TITLE OF INVENTION: FATTY ACID ELONGASE 3-KETOACYL COA
; TITLE OF INVENTION: SYNTHASE POLYPEPTIDES
; FILE REFERENCE: 07148-108001
; CURRENT APPLICATION NUMBER: US/09/877,476
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 60/210,326
; PRIOR FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 506
; TYPE: PRT
; ORGANISM: Artificial Sequence

```



```

; FEATURE:
; OTHER INFORMATION: 5' 114 amino acids from A. thaliana PAE1 (SEQ ID
; OTHER INFORMATION: NO:2) and 3' 392 amino acids from B. napus
; OTHER INFORMATION: elongase KCS (SEQ ID NO:4) having mutations at
; OTHER INFORMATION: positions 91, 92 and 307; designated At114 L91C
; OTHER INFORMATION: K92R G307D; hypothetical
US-09-877-476-20

```

```

Query Match          38.2%; Score 47; DB 9; Length 506;
Best Local Similarity 71.4%; Pred. No. 1e+02;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 9 PPLSTFLLGIASRL 22
    |||: ||| |||
Db 23 PPLTAFLAGKASRL 36

```

```

RESULT 15
US-09-877-476-22
; Sequence 22, Application US/09877476
; Patent No. US20020049994A1
; GENERAL INFORMATION:
; APPLICANT: Jaworski, Jan G.
; APPLICANT: Blacklock, Brenda J.
; TITLE OF INVENTION: FATTY ACID ELONGASE 3-KETOACYL COA
; TITLE OF INVENTION: SYNTHASE POLYPEPTIDES
; FILE REFERENCE: 07148-108001
; CURRENT APPLICATION NUMBER: US/09/877,476
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 60/210,326
; PRIOR FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 506
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: 5' 114 amino acids from A. thaliana PAE1 (SEQ ID
; OTHER INFORMATION: NO:2) and 3' 392 amino acids from B. napus
; OTHER INFORMATION: elongase KCS (SEQ ID NO:4) having mutations at
; OTHER INFORMATION: positions 92 and 307; designated At114 K92R G307D;
; OTHER INFORMATION: hypothetical
US-09-877-476-22

```

```

Query Match          38.2%; Score 47; DB 9; Length 506;
Best Local Similarity 71.4%; Pred. No. 1e+02;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 9 PPLSTFLLGIASRL 22
    |||: ||| |||
Db 23 PPLTAFLAGKASRL 36

```

Search completed: November 11, 2004, 01:28:23
Job time : 27.9809 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 12:29:32 ; Search time 5.79275 Seconds
(without alignments)
431.857 Million cell updates/sec

Title: US-10-092-750-42

Perfect score: 123
Sequence: 1 GVSEAEGTFFLSTFLGIASRLRSVA 26

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79:*

1: Pir1.*

2: Pir2.*

3: Pir3.*

4: Pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50.5	41.1	209	B90497	conserved hypot
2	48	39.0	154	H72769	hypothetical prote
3	47.5	38.6	1467	T48152	hypothetical prote
4	47	38.2	275	A56641	probable membrane
5	47	38.2	487	G83827	stage V sporulatio
6	47	38.2	506	T05272	fatty acid elongas
7	47	38.2	513	T43434	translation elonga
8	47	38.2	1047	T41343	probable translati
9	46.5	37.8	153	AB1415	hypothetical prote
10	46.5	37.8	478	T45651	hypothetical prote
11	46.5	37.8	839	AC2637	large atp-dependan
12	46.5	37.8	848	B97419	probable ATP-depen
13	46	37.4	113	S59116	programmed cell de
14	46	37.4	126	AE0326	probable membrane
15	46	37.4	254	C70356	flagellar motor pr
16	46	37.4	358	H90052	conserved hypot
17	46	37.4	378	G71290	flagellar biosynth
18	46	37.4	485	G85943	probable transport
19	46	37.4	485	B65072	probable transport
20	46	37.4	485	C91098	probable transport
21	46	37.4	581	C71078	aldehyde-ferradoxi
22	46	37.4	1469	H96522	probable ABC trans
23	45.5	37.0	555	T38479	myb-like DNA-bindi
24	45.5	37.0	853	A71339	probable outer mem
25	45	36.6	113	B54437	apoptotic cell dea
26	45	36.6	123	AC3331	integral membrane
27	45	36.6	209	D72029	endonuclease III C
28	45	36.6	209	C86595	endonuclease III I
29	45	36.6	466	T26458	hypothetical prote

30	45	36.6	665	2	T33743	hypothetical prote
31	45	36.6	672	2	G88651	protein B0212.3 li
32	45	36.6	759	1	C64345	replication licens
33	45	36.6	993	2	A96750	hypothetical prote
34	45	36.6	1436	2	B70520	probable PPE prote
35	44	35.8	197	2	S51464	probable membrane
36	44	35.8	252	2	G72298	transcription regu
37	44	35.8	349	2	I55601	Na/taurocholate co
38	44	35.8	446	2	F70302	transporter (Pho87
39	44	35.8	449	2	T29365	hypothetical prote
40	44	35.8	460	2	C82399	conserved hypot
41	43.5	35.4	153	2	A11790	hypothetical prote
42	43.5	35.4	253	2	T35398	probable ABC trans
43	43.5	35.4	500	2	T22068	hypothetical prote
44	43.5	35.4	618	2	E82691	RNA polymerase sig
45	43	35.0	113	2	A54437	apoptotic cell dea

ALIGNMENTS

RESULT 1

B90497

conserved hypothetical protein [imported] - Sulfolobus solfataricus

C;Species: Sulfolobus solfataricus

C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004

C;Accession: B90497

R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-

arrrett, R.A.; Ragan, M.A.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P

submitted to GenBank, April 2001

A;Description: Sulfolobus solfataricus complete genome.

A;Reference number: A99139

A;Accession: B90497

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-209 <KUR>

A;Cross-references: UNIPROT:Q97U88; GB:AE006641; NID:gi13816556; PIDN:AAK43233.1; GSPDB:G

C;Genetics:

A;Gene: SSO3132

C;Superfamily: glyoxalase

Query Match 41.1%; Score 50.5; DB 2; Length 209;
Best Local Similarity 41.9%; Pred. No. 2, 1;
Matches 13; Conservative 5; Mismatches 8; Indels 5; Gaps 1;

QY 1 GVSEAEGTFFLSTFL-----GIASRLRSVA 26

|||: : ||||| |||: | :|

Db 62 GVNQIKRDPSPSTFLINEXDLGLKXASSMA 92

RESULT 2

H72769

hypothetical protein APE0147 - Aeropyrum pernix (strain K1)

C;Species: Aeropyrum pernix

C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004

C;Accession: H72769

R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Ki

DNA Res. 6, 83-101, 1999

A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr

A;Reference number: A72450; MUID:99310339; PMID:10382966

A;Accession: H72769

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-154 <KAW>

A;Cross-references: UNIPROT:Q9YFV3; DDBJ:AP000058; NID:95103388; PIDN:BAA79058.1; PID:dl

A;Experimental source: strain K1

C;Genetics:

A;Gene: APE0147

C;Superfamily: conserved hypothetical protein MJ0281

Query Match 39.0%; Score 48; DB 2; Length 154;

Best Local Similarity 42.3%; Pred. No. 3.9;
Matches 11; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 1 GVSEAGTFFLSTFLGLIASRLRVA 26
DB 53 GVDADGVIYGFNFVGLASTALGVA 78

RESULT 3
T48162
hypothetical protein T1008.110 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
C:Accession: T48162
R:Bevan, M.; Pohl, T.; Weizenegger, T.; Bancroft, I.; Mewes, H.W.; Lemcke, K.; Mayer, K.
submitted to the Protein Sequence Database, March 2000
A:Reference number: Z24486
A:Accession: T48162
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-1467 <BEV>
A:Cross-references: UNIPROT:Q9M033; EMBL:AL161746
A:Experimental source: cultivar Columbia; BAC clone T1008
C:Genetics:
A:Map position: 5
A:Introns: 58/2; 121/3; 152/3; 178/1; 234/2; 291/3; 331/3; 516/3; 698/3; 721/3; 760/3; 82/3
A:Note: T1008.110

Query Match 38.6%; Score 47.5; DB 2; Length 1467;
Best Local Similarity 58.3%; Pred. No. 48;
Matches 14; Conservative 2; Mismatches 7; Indels 1; Gaps 1;

QY 2 VSEAGTFFLSTFLGLIASRLRS 24
DB 1146 VLQAGAFPALSDFILEILSLRVS 1169

RESULT 4
A56641
probable membrane transport protein - Clostridium perfringens
C:Species: Clostridium perfringens
C:Date: 11-Aug-1995 #sequence_revision 11-Aug-1995 #text_change 09-Jul-2004
C:Accession: A56641
R:Holck, A.L.; Blom, H.
DNA Seq. 3, 191-194, 1992
A:Title: The nucleotide sequence of a putative membrane transport gene from Clostridium
A:Reference number: A56641; MUID:93113001; PMID:1472712
A:Accession: A56641
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-275 <HO>
A:Cross-references: UNIPROT:Q06111; GB:X66092; NID:9296355; PIDN:CRA46887.1; PID:9296356
A:Note: nucleotide sequence not given; conceptual translation not complete
C:Superfamily: maltose transport protein malG

Query Match 38.2%; Score 47; DB 2; Length 275;
Best Local Similarity 38.5%; Pred. No. 10;
Matches 10; Conservative 3; Mismatches 13; Indels 0; Gaps 0;

QY 1 GVSEAGTFFLSTFLGLIASRLRVA 26
DB 146 GIVSAFGTLLRQFPMGLPKELEAA 171

RESULT 5
G83827
stage V sporulation protein AF BHI423 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C:Accession: G83827
R:Takami, H.; Nakasono, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: AB3650; MUID:20512582; PMID:11056132
A:Accession: G83827
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-487 <STO>
A:Cross-references: UNIPROT:Q9KCZ6; GB:AP001512; GB:BA000004; NID:gi0174030; PIDN:BA00514
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BHI423

Query Match 38.2%; Score 47; DB 2; Length 487;
Best Local Similarity 50.0%; Pred. No. 18;
Matches 11; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 5 AEGTFFLSTFLGLIASRLRVA 26
DB 394 AVGTATPSYELGLANLIRVA 415

RESULT 6
T05272
fatty acid elongase 1 - Arabidopsis thaliana
N:Alternate names: protein T4L20.100
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: T05272
R:Bevan, M.; Terry, N.; Ardiles, W.; Buysshaert, C.; Dasseville, R.; De Clerck, R.; De
ewes, H.W.; Mayer, K.F.X.; Schueller, C.
submitted to the Protein Sequence Database, September 1998
A:Reference number: Z15406
A:Accession: T05272
A:Molecule type: DNA
A:Residues: 1-506 <BEV>
A:Cross-references: UNIPROT:Q38860; EMBL:AL023094
A:Experimental source: cultivar Columbia; BAC clone T4L20
C:Genetics:
A:Map position: 4
A:Note: T4L20.100
C:Superfamily: very-long-chain 3-ketoacyl-CoA synthase

Query Match 38.2%; Score 47; DB 2; Length 506;
Best Local Similarity 71.4%; Pred. No. 19;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 9 FPLSTFLGLIASRL 22
DB 23 FLTAFLAGKASRL 36

RESULT 7
T43434
translation elongation factor eRF-3 - fission yeast (Schizosaccharomyces pombe) (fragment
C:Species: Schizosaccharomyces pombe
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
C:Accession: T43434
R:Uritani, M.; Shoumura, Y.; Yamada, S.
Biosci. Biotechnol. Biochem. 63, 769-772, 1999
A:Title: Detection and analysis of translation elongation factor 3 genes from various yeast
A:Reference number: Z22510; MUID:99290044; PMID:10361693
A:Accession: T43434
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-513 <URI>
A:Cross-references: UNIPROT:Q94489; EMBL:AB018538; NID:93776153; PIDN:BAA33896.1; PID:937
A:Experimental source: strain JY741
C:Genetics:
A:Gene: RF-3
C:Superfamily: translation elongation factor 3; ATP-binding cassette homology

Query Match 38.2%; Score 47; DB 2; Length 513;
Best Local Similarity 40.0%; Pred. No. 20;
Matches 10; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 1 GVSEAGTFFPLSTFLGASRLRSV 25
 |||||
 Db 199 GASEMEKFPFPGFLEGVKTQKRAI 223
 |||||

RESULT 8
 T41343
 probable translation elongation factor EF-3 - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 C:Accession: T41343
 R:Lyne, M.; Rajandream, M.A.; Barrrell, B.G.; Volckaert, G.
 submitted to the EMBL Data Library, January 1999
 A:Reference number: Z21970
 A:Accession: T41343
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1047 <LYN>
 A:Cross-references: UNIPROT:O94489; EMBL:AL035076; PIDN:CAA22654.1; GSPDB:GN000066; SPDB:
 A:Experimental source: strain 972h-; cosmid c417
 C:Genetics:
 A:Gene: SPDB:SPCC417.08
 A:Map position: 1
 C:Superfamily: translation elongation factor 3; ATP-binding cassette homology

Query Match 38.2%; Score 47; DB 2; Length 1047;
 Best Local Similarity 40.0%; Pred. No. 41;
 Matches 10; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 1 GVSEAGTFFPLSTFLGASRLRSV 25
 |||||
 Db 648 GASEMEKFPFPGFLEGVKTQKRAI 672
 |||||

RESULT 9
 AB1415
 hypothetical protein lmo2723 [imported] - Listeria monocytogenes (strain EGD-e)
 C:Species: Listeria monocytogenes
 C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
 C:Accession: AB1415
 R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
 ; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
 D.; Jones, L.M.; Karst, U.
 Science 294, 849-852, 2001
 A:Authors: Kref, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
 ck, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Varquez-Boland, J.A.; Voss, H.; Wehlend,
 A:Title: Comparative genomics of Listeria species.
 A:Reference number: AB1077; MUID:21537279; PMID:11679669
 A:Accession: AB1415
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-153 <GLA>
 A:Cross-references: UNIPROT:Q8V3V8; GB:NC_003210; PIDN:CAD00936.1; PID:G16412223; GSPDB:
 A:Experimental source: strain EGD-e
 C:Genetics:
 A:Gene: lmo2723

Query Match 37.8%; Score 46.5; DB 2; Length 153;
 Best Local Similarity 35.7%; Pred. No. 6.7;
 Matches 15; Conservative 3; Mismatches 5; Indels 19; Gaps 2;

QY 4 BAEGTFFPLSTFLGASRLRSV 26
 :|||||
 Db 34 DMEGTFFLGAPEKDVLLGVAFPEKSTVIMNPQYRIGVA 75
 :|||||

RESULT 10
 T45661
 hypothetical protein F13112.300 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
 C:Accession: T45661

R;Choisne, N.; Robert, C.; Brottier, P.; Wincker, P.; Cattolico, L.; Artiguenave, F.; Sai
 submitted to the Protein Sequence Database, November 1999
 A:Reference number: Z23010
 A:Accession: T45661
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-478 <CHO>
 A:Cross-references: UNIPROT:Q9SD48; EMBL:AL1133292
 A:Experimental source: cultivar Columbia; BAC clone F13112
 C:Genetics:
 A:Map position: 3
 A:Introns: 255/3
 A:Note: F13112.300

Query Match 37.8%; Score 46.5; DB 2; Length 478;
 Best Local Similarity 50.0%; Pred. No. 22;
 Matches 15; Conservative 1; Mismatches 9; Indels 5; Gaps 1;

QY 2 VSEAGTFFPLSTFLGASRLRSV 26
 |||||
 Db 285 VSSSESTFFLILSAKRLRQGIKFLRSDA 314
 |||||

RESULT 11
 AC2637
 large atp-dependant helicase-related protein [imported] - Agrobacterium tumefaciens (str.
 C:Species: Agrobacterium tumefaciens
 C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
 C:Accession: AC2637
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell,
 ; Karp, P.; Romero, P.; Zhang, S.
 Science 294, 2317-2323, 2001
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, I.
 ster, B.W.
 A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
 A:Reference number: AB2577; MUID:21608550; PMID:11743193
 A:Accession: AC2637
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-839 <XUR>
 A:Cross-references: UNIPROT:Q8UI06; GB:AE008688; PIDN:AAL41513.1; PID:G17738842; GSPDB:G
 A:Experimental source: strain C58 (Duport)
 C:Genetics:
 A:Gene: lhr
 A:Map position: circular chromosome

Query Match 37.8%; Score 46.5; DB 2; Length 839;
 Best Local Similarity 52.2%; Pred. No. 39;
 Matches 12; Conservative 4; Mismatches 4; Indels 3; Gaps 1;

QY 3 SEAGTFFPLSTFLGASRLRSV 25
 |||||
 Db 570 SYAGKFFPLSTYL--ADQVRSM 589
 |||||

RESULT 12
 B97419
 probable ATP-dependent helicase mj0294 [imported] - Agrobacterium tumefaciens (strain C58
 C:Species: Agrobacterium tumefaciens
 C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
 C:Accession: B97419
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,
 A.; Liu, F.; Wollam, C.; Allinger, D.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
 Science 294, 2323-2328, 2001
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
 A:Reference number: A97359; MUID:21608551; PMID:11743194
 A:Accession: B97419
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-848 <XUR>
 A:Cross-references: UNIPROT:Q8UI06; GB:AE007869; PIDN:AAK86307.1; PID:G15155423; GSPDB:G
 C:Genetics:

```

A:Gene: AGR_C_874
A:Map position: circular chromosome

Query Match      37.8%; Score 46.5; DB 2; Length 848;
Best Local Similarity 52.2%; Pred. No. 39;
Matches 12; Conservative 4; Mismatches 4; Indels 3; Gaps 1;

QY 3 SEAGTFLPLSTFLGLGASRLRSV 25
    |||||
Db 579 SYAGKFLSTYL---ADQVRSM 598
    |||||

RESULT 13
S59116
programmed cell death suppressor dad-1 - Caenorhabditis elegans
N:Alternate names: gene defender against cell death protein
C:Species: Caenorhabditis elegans
C>Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 09-Jul-2004
C:Accession: S59116; T32742; S57513
R:Submitter: A.; Hozak, R.R.; Nakashima, T.; Nishimoto, T.; Rothman, J.H.
EMBO J. 14, 4434-4441, 1995
A:Title: dad-1, an endogenous programmed cell death suppressor in Caenorhabditis elegans
A:Reference number: S59116; MUID:96003623; PMID:7556086
A:Accession: S59116
A:Molecule type: mRNA
A:Residues: 1-113 <SUG>
A:Cross-references: UNIPROT:P52872; EMBL:X89080; NID:9887395; PIDN:CAA61451.1; PID:98873
R:Greco, T.; Elliott, G.; Keppeler, D.
submitted to the EMBL Data Library, December 1997
A:Description: The sequence of C. elegans cosmid F57B10.
A:Reference number: Z21219
A:Accession: T32742
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-113 <GRE>
A:Cross-references: EMBL:AF039713; PIDN:AAB96727.1; GSPDB:GN00019; CESP:F57B10.10
A:Experimental source: strain Bristol N2; clone F57B10
C:Genetics:
A:Gene: dad-1; CESP:F57B10.10
A:Map position: 1
A:Introns: 71/1
C:Superfamily: apoptotic cell death regulator DAD1
C:Keywords: apoptosis

Query Match      37.4%; Score 46; DB 2; Length 113;
Best Local Similarity 50.0%; Pred. No. 5.9;
Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 7 GTFFPLSTFLGLGASRLRS 24
    |||||
Db 49 GTFFPNSFLSGFISTVTS 66
    |||||

RESULT 14
AE0326
probable membrane protein YPO2677 [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C:Accession: AE0326
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
Hill, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AE0001; MUID:21470413; PMID:11586360
A:Accession: AE0326
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-126 <KUR>
A:Cross-references: UNIPROT:Q8ZDB1; GB:AL590842; PIDN:CAC92916.1; PID:g15980656; GSPDB:G
C:Genetics:
A:Gene: YPO2677
C:Superfamily: hypothetical protein MUI523

```

```

Query Match      37.4%; Score 46; DB 2; Length 126;
Best Local Similarity 57.1%; Pred. No. 6.6;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 6 EGTFFPLSTFLGLGIA 19
    |||||
Db 32 KGNFPLGLTFLINIS 45
    |||||

RESULT 15
G70386
flagellar motor protein MotA - Aquifex aeolicus
C:Species: Aquifex aeolicus
C>Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004
C:Accession: G70386
R:Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Over
V. Nature 392, 353-358, 1998
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; MUID:98196666; PMID:9537320
A:Accession: G70386
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-254 <AQP>
A:Cross-references: UNIPROT:O67122; GB:AE000718; NID:g2983504; PIDN:AAC07083.1; PID:g2983
A:Experimental source: strain VFS
C:Genetics:
A:Gene: motA
C:Superfamily: flagellar motor rotation protein

Query Match      37.4%; Score 46; DB 2; Length 254;
Best Local Similarity 44.4%; Pred. No. 14;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 GVSAGETFFPLSTFLGLI 18
    |||||
Db 42 GNAAMGAPFLKDFIRGV 59
    |||||

Search completed: November 10, 2004, 13:40:49
Job time : 5.79275 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 12:27:34 ; Search time 31.7478 Seconds
(without alignments)
471.205 Million cell updates/sec

Title: US-10-092-750-42
Perfect score: 123
Sequence: 1 GVSEAGTFPLSTFLLGASRLRSA 26

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

```
Database :      UniProt_02:*
1:  uniprot_sprot:*
2:  uniprot_trembl:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query		DB	ID	Description
	Score	Match Length			
1	52	42.3	402	2	Q69f85 uncultured
2	52	42.3	402	2	AAR38338
3	52	42.3	429	2	Q69GC7 uncultured
4	52	42.3	429	2	AAR37935 uncultured
5	51	41.5	705	2	Q9ZNM3 vibrio orie
6	50.5	41.1	209	2	Q97U88
7	48	39.0	154	2	Q97U88 sulfolobus
8	47.5	38.6	153	2	Q9YfV3 aeropyrum p
9	47.5	38.6	153	2	Q71W41 listeria mo
10	47.5	38.6	1467	2	AAT05475 listeria
11	47	38.2	275	2	Q9M033 arabidopsis
12	47	38.2	487	2	Q06111 clostridium
13	47	38.2	491	2	Q9KcZ6 bacillus ha
14	47	38.2	506	2	Q8W5D1 oryza sativ
15	47	38.2	532	2	Q38860 arabidopsis
16	47	38.2	740	2	Q6BTK1 debaryomyce
17	47	38.2	1042	2	Q87378 acetobacter
18	47	38.2	1047	1	Q81m18 oryza sativ
19	47	38.2	1056	2	Q94489 schizosacch
20	46.5	37.8	153	2	Q759Z6 neurospora
21	46.5	37.8	480	2	Q873V8 listeria mo
22	46.5	37.8	480	2	Q81703 arabidopsis
23	46.5	37.8	839	2	Q98d48 arabidopsis
24	46.5	37.8	848	2	Q8ui06 agrobacteri
25	46	37.4	113	1	Q7D1C7 agrobacteri
26	46	37.4	126	1	P52872 caenorhabdi
27	46	37.4	254	1	Q8zdbl versinia pe
28	46	37.4	277	2	Q67122 aquifex ae
29	46	37.4	277	2	Q73A81 bacillus ce
30	46	37.4	358	2	AAS40825 bacillus
31	46	37.4	358	2	Q86119 dictyosteli
	46	37.4	358	2	Q99rd8 staphylococ

32	46	37.4	358	2	Q7A3K3	Q7A3K3	staphylococ
33	46	37.4	376	1	FUHB_TREPA	Q83710	treponema p
34	46	37.4	397	2	Q6G6H9	Q6G6H9	staphylococ
35	46	37.4	397	2	Q8NUV1	Q8NUV1	staphylococ
36	46	37.4	485	1	YQFO_ECOLI	Q46815	escherichia
37	46	37.4	581	2	Q58642	Q58642	pyrococcus
38	46	37.4	605	2	Q82LS8	Q82LS8	streptomyce
39	46	37.4	624	1	RPOD_XANCP	Q8P4H2	xanthomonas
40	46	37.4	815	2	Q72CJ3	Q72CJ3	desulfovibr
41	46	37.4	815	2	AS95S950	AS95S950	desulfovi
42	46	37.4	1469	2	Q9XIE2	Q9XIE2	arabidopsis
43	45.5	37.0	1112	2	Q6TEC7	Q6TEC7	anopheles g
44	45.5	37.0	112	2	AAQ9A040	AAQ9A040	anopheles
45	45.5	37.0	384	2	Q9K3H9	Q9K3H9	streptomyce

ALIGNMENTS

```

RESULT 1
ID ID Q6SF85 PRELIMINARY; PRT; 402 AA.
AC Q6SF85;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Drug resistance transporter, Bcr/CflA family protein.
GN ORFNames=EBAC000-69B03.87;
OS uncultured bacterium 581.
OC Bacteria; environmental samples.
OX NCBI_TaxID=257401;
RN [1]
RP SEQUENCE FROM N.A.
RA DeLong E.F.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
RL [2]
RP SEQUENCE FROM N.A.
RA Heidelberg J.F., Eisen J.A., Nelson W.C., DeLong E.F.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
RL EMBL; AY45648; AAR38338.1; -;
DR InterPro; IPR004812; Eflux_Bcr_CflA.
DR InterPro; IPR007114; MFS.
DR TIGRPFAM; TIGR00710; eflux_Bcr_CflA; 1.
DR PROSITE; PS00850; MFS; 1.
SQ SEQUENCE 402 AA; 42022 MW; 9655B3DAD962D69F CRC64;

Query Watch 42.3%; Score 52; DB 2; Length 40
Best Local Similarity 52.6%; Pred No; 16;
Matches 10; Conservative 4; Mismatches 5; Indels

QY 1 GVSEAGTFPLSTFLGLGIA 19
- - - - -
DB 36 GATASDQQLTSLTMLGIA 54

RESULT 2
AAR38338
ID AAR38338 PRELIMINARY; PRT; 402 AA.
AC AAR38338;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Drug resistance transporter, Bcr/CflA family protein.
GN EBAC000-69B03.87.
OS uncultured bacterium 581.
OC Bacteria; environmental samples.
OX NCBI_TaxID=257401;
RN [1]
RP SEQUENCE FROM N.A.
RA DeLong E.F.;
RL "Monterey Bay Coastal Ocean Microbial Observatory environ-
RL sequencing."
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.

```

```

RN [2]
RP SEQUENCE FROM N.A.
RA Heidelberg J.F., Eisen J.A., Nelson W.C., DeLong E.F.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY458648; AAR37935.1; -.
SQ SEQUENCE 402 AA; 42022 MW; 9655B3DAD962D69F CRC64;

Query Match      42.3%; Score 52; DB 2; Length 402;
Best Local Similarity 52.6%; Pred. No. 16;
Matches 10; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 GVSEAGTFFPLSTFLGLIA 19
Db 36 GATASDGGTLSTFLMGLIA 54

RESULT 3
QSGC7
ID AAR37935 PRELIMINARY; PRT; 429 AA.
AC Q6SGC7;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Drug resistance transporter, Bcr/CflA family.
GN ORFNames=EBAC000-47H08.14;
OS uncultured bacterium 561;
OC Bacteria; environmental samples.
OX NCBI_TaxID=257396;
RN [1]_TaxID=257396;
RP SEQUENCE FROM N.A.
RA DeLong E.F.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Heidelberg J.F., Eisen J.A., Nelson W.C., DeLong E.F.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY458643; AAR37935.1; -.
DR InterPro; IPR004812; Eflux_Bcr_CflA.
DR InterPro; IPR007114; MFS.
DR TIGRPFAMs; TIGR00710; eflux_Bcr_CflA; 1.
DR PROSITE; PS00850; MFS; 1.
SQ SEQUENCE 429 AA; 44843 MW; E763DC581B810193 CRC64;

Query Match      42.3%; Score 52; DB 2; Length 429;
Best Local Similarity 52.6%; Pred. No. 17;
Matches 10; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 GVSEAGTFFPLSTFLGLIA 19
Db 53 GATASDGGTLSTFLMGLIA 81

RESULT 4
AAR37935
ID AAR37935 PRELIMINARY; PRT; 429 AA.
AC AAR37935;
DT 02-MAR-2004 (TReMBLrel. 27, Created)
DT 02-MAR-2004 (TReMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TReMBLrel. 27, Last annotation update)
DE Drug resistance transporter, Bcr/CflA family.
GN EBAC000-47H08.14.
OS uncultured bacterium 561.
OC Bacteria; environmental samples.
OX NCBI_TaxID=257396;
RN [1]_TaxID=257396;
RP SEQUENCE FROM N.A.
RA DeLong E.F.;
RL Monterey Bay Coastal Ocean Microbial Observatory environmental clone sequencing;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Heidelberg J.F., Eisen J.A., Nelson W.C., DeLong E.F.;

```

```

RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY458643; AAR37935.1; -.
SQ SEQUENCE 429 AA; 44843 MW; E763DC581B810193 CRC64;

Query Match      42.3%; Score 52; DB 2; Length 429;
Best Local Similarity 52.6%; Pred. No. 17;
Matches 10; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 GVSEAGTFFPLSTFLGLIA 19
Db 63 GATASDGGTLSTFLMGLIA 81

RESULT 5
Q9ZNM3
ID Q9ZNM3 PRELIMINARY; PRT; 705 AA.
AC Q9ZNM3;
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE IutA.
GN Name=iutA;
OS Vibrio orientalis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=28175;
RN [1]
RP SEQUENCE FROM N.A.
RA Murakami K., Fuse H., Takimura O., Inoue H., Yamaoka Y.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Outer membrane (By similarity).
DR EMBL; AB010890; BAA74703.1; -.
DR GO; GO:0019867; B:outer membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000531; TonB_receptor.
DR InterPro; IPR010917; TonB_recept_C.
DR InterPro; IPR010105; TonB_sdp_recept.
DR Pfam; PF00593; TonB_dep_Rec; 1.
DR TIGRPFAMs; TIGR01783; TonB-siderophor; 1.
DR PROSITE; PS01156; TONB_DEPENDENT_REC_2; 1.
DR Membrane; Outer membrane; Receptor; TonB box.
RW
SQ SEQUENCE 705 AA; 77907 MW; 440CB4C074E44212 CRC64;

Query Match      41.5%; Score 51; DB 2; Length 705;
Best Local Similarity 60.0%; Pred. No. 39;
Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 6 EGTFFPLSTFLGLIAS 20
Db 5 KGSFFLSTIALGVAA 19

RESULT 6
Q97U88
ID Q97U88 PRELIMINARY; PRT; 209 AA.
AC Q97U88;
DT 01-OCT-2001 (TReMBLrel. 18, Created)
DT 01-OCT-2001 (TReMBLrel. 18, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=SSO31132;
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=2287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=21332296; PubMed=11427726;
RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA Awayez M.J., Chan-Weiher C.C.-Y., Clausen I.G., Curtis B.A.,

```



```

RA De Moors A., Erauso G., Fletcher C., Gordon P.M.K.,
RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.,
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.",
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
DR EMBL; AE006903; AA043233.1; -
DR PIR; B90497; B90497.
DR InterPro; IPR001279; Blackmase-like.
DR Pfam; PF00753; Lactamase_B; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 209 AA; 23000 MW; AA863D463AB056A6 CRC64;

Query Match 41.1%; Score 50.5; DB 2; Length 209;
Best Local Similarity 41.9%; Pred. No. 15;
Matches 13; Conservative 5; Mismatches 8; Indels 5; Gaps 1;

Oy 1 GVSAEGTFFPLSTFL-----LGIASRLRVA 26
|||: ||| ||| ||| ||| ||| ||| ||| |||
Db 62 GVNQIKRDFPSSTFLINEKDLGLKRASSMA 92

RESULT 7
O9YFV3 PRELIMINARY; PRT; 154 AA.
AC O9YFV3;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein APE0147.
GN OrderedLocusNames=APE0147;
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OC Desulfurococaceae; Aeropyrum.
CX NCBI_TaxID=56636;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RX MEDLINE=99310339; PubMed=10382966;
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101(1999).
DR EMBL; AP000058; BAA79058.1; -
DR PIR; H72769; H72769.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 154 AA; 17209 MW; 1D0C1270DAD8BB7 CRC64;

Query Match 39.0%; Score 48; DB 2; Length 154;
Best Local Similarity 42.3%; Pred. No. 27;
Matches 11; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

Oy 1 GVSAEGTFFPLSTFLGLIASRLRVA 26
|||: ||| ||| ||| ||| ||| ||| |||
Db 53 GVDADGYIPGFNFVGLASTALGVA 78

RESULT 8
Q71W41 PRELIMINARY; PRT; 153 AA.
ID Q71W41;
AC Q71W41;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Acetyltransferase, GNAT family.
GN OrderedLocusNames=LMOF2365_2710;
OS Listeria monocytogenes (serotype 4b / strain F2365).

```

```

OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=265669;
RN [1];
RP SEQUENCE FROM N.A.
RX PubMed=15115801; DOI=10.1093/nar/gkh562;
RA Nelson K.E., Fouts D.E., Mongodin E.F., Gill S.R., Paulsen I.T.,
RA Kolonay J.F., Raeko D.A., Angiuoli S.V., Nierman W.C., Beanan M.J.,
RA Peterson J.D., White O., Nelson W.C., Durkin A.S., Madupu R.,
RA Brinkac L.M., Daugherty S.C., Dodson R.J., Durkin A.S., Madupu R.,
RA Haft D.H., Selengut J., Van Aken S.E., Khouri H.M., Fedorova N.,
RA Forbarger H.A., Tran B., Kathariou S., Wonderling L.D., Unlich G.A.,
RA Bayles D.O., Luchansky J.B., Fraser C.M.;
RT "Whole genome comparisons of serotype 4b and 1/2a strains of the food-
RT borne pathogen Listeria monocytogenes reveal new insights into the
RT core genome components of this species.";
RL Nucleic Acids Res. 32:2386-2395(2004).
DR EMBL; AE017331; AA05475.1; -
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR00182; GCNSacetyl_trans.
DR Pfam; PF00583; Acetyltransf_1; 1.
KW Complete proteome; Transferase.
SQ SEQUENCE 153 AA; 17480 MW; A4D231EC38ED5CCF CRC64;

Query Match 38.6%; Score 47.5; DB 2; Length 153;
Best Local Similarity 38.1%; Pred. No. 32;
Matches 16; Conservative 2; Mismatches 5; Indels 19; Gaps 2;

Oy 4 EAEGTFFPLSTF-----LLGIAS-----RLRSVA 26
: ||| ||| ||| ||| ||| ||| ||| |||
Db 34 DMEGTHLGAFEKDVLLGIASFPYKSTVIMNPAQYRIGVA 75

RESULT 9
AA05475 PRELIMINARY; PRT; 153 AA.
ID AA05475;
AC AA05475;
DT 10-MAY-2004 (TrEMBLrel. 27, Created)
DT 10-MAY-2004 (TrEMBLrel. 27, Last sequence update)
DT 10-MAY-2004 (TrEMBLrel. 27, Last annotation update)
DE Acetyltransferase, GNAT family.
GN LMOF2365_2710.
OS Listeria monocytogenes str. 4b F2365.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria;
OC Listeria monocytogenes.
CX NCBI_TaxID=265669;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=4b F2365;
RX PubMed=15115801;
RA Nelson K.E., Fouts D.E., Mongodin E.F., Ravel J., DeBoy R.T.,
RA Kolonay J.F., Rasko D.A., Angiuoli S.V., Gill S.R., Paulsen I.T.,
RA Peterson J., White O., Nelson W.C., Nierman W., Beanan M.J.,
RA Brinkac L.M., Daugherty S.C., Dodson R.J., Durkin A.S., Madupu R.,
RA Haft D.H., Selengut J., Van Aken S., Khouri H., Fedorova N.,
RA Forbarger H., Tran B., Kathariou S., Wonderling L.D., Unlich G.A.,
RA Bayles D.O., Luchansky J.B., Fraser C.M.;
RT "Whole genome comparisons of serotype 4b and 1/2a strains of the food-
RT borne pathogen Listeria monocytogenes reveal new insights into the
RT core genome components of this species.";
RL Nucleic Acids Res. 32:2386-2395(2004).
DR EMBL; AE017331; AA05475.1; -
KW Transferase.
SQ SEQUENCE 153 AA; 17480 MW; A4D231EC38ED5CCF CRC64;

Query Match 38.6%; Score 47.5; DB 2; Length 153;
Best Local Similarity 38.1%; Pred. No. 32;
Matches 16; Conservative 2; Mismatches 5; Indels 19; Gaps 2;

Oy 4 EAEGTFFPLSTF-----LLGIAS-----RLRSVA 26
: ||| ||| ||| ||| ||| ||| ||| |||
Db 34 DMEGTHLGAFEKDVLLGIASFPYKSTVIMNPAQYRIGVA 75

```

```

RESULT 10
Q9M033 ID Q9M033 PRELIMINARY; PRT; 1467 AA.
AC Q9M033;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein T1008_110.
GN Name=T1008_110;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Pohl T., Weizenegger T., Bancroft I., Mewes H.W., Rudd S.,
RA Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL161746; CAB81923.1; -
DR PIR; T48162; T48162.
DR InterPro; IPR008938; ARM.
KW Hypothetical protein.
SQ SEQUENCE 1467 AA; 159477 MW; 414CA3C4DD9EA705 CRC64;

Query Match 38.6%; Score 47.5; DB 2; Length 1467;
Best Local Similarity 58.3%; Pred. No. 2.9e+02;
Matches 14; Conservative 2; Mismatches 7; Indels 1; Gaps 1;

QY 2 VSEAGTTP-ISTFLGLIASRLRS 24
DB 1146 VLQAIGAPFALSDFILELSRLVS 1169

RESULT 11
Q06111 ID Q06111 PRELIMINARY; PRT; 275 AA.
AC Q06111;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ORF for putative membrane transport protein.
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1502;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=NCTC8239;
RX MEDLINE=93113001; PubMed=1472712;
RA Holck A.L., Blom H.;
RT "The nucleotide sequence of a putative membrane transport gene from
RT Clostridium perfringens."
RL DNA Seq. 3:191-194 (1992).
CC -!- FUNCTION: Part of a binding-protein-dependent transport system.
CC Probably responsible for the translocation of the substrate across
CC the membrane (By similarity)
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the binding-protein-dependent transport
CC system permease family.
DR EMBL; X66092; CAA46887.1; -
DR PIR; A56641; A56641.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000515; BPD_transp.
DR Pfam; PF00528; BPD_transp_1; 1.
DR PROSITE; PS50928; ABC_TM1; 1.
KW Transmembrane; Transport.

```

```

SQ SEQUENCE 275 AA; 30699 MW; CC6B951258772772 CRC64;

Query Match 38.2%; Score 47; DB 2; Length 275;
Best Local Similarity 38.5%; Pred. No. 68;
Matches 10; Conservative 3; Mismatches 13; Indels 0; Gaps 0;

QY 1 GVSEAGTTP-ISTFLGLIASRLRSVA 26
DB 146 GIVSAPGTFLRQFFMGLPKLEEEAA 171

RESULT 12
Q9KCZ6 ID Q9KCZ6 PRELIMINARY; PRT; 487 AA.
AC Q9KCZ6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Stage V sporulation protein AF.
GN Name=BHI423;
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=C-125;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis."
RL Nucleic Acids Res. 28:4317-4331 (2000).
DR EMBL; AF001512; BAB05142.1; -
DR PIR; G83827; G83827.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0009847; P:spore germination; IEA.
DR InterPro; IPR004995; GeraA.
DR Pfam; PF03323; GeraA; 1.
SQ SEQUENCE 487 AA; 54412 MW; FD9A224B1A1CB149 CRC64;

Query Match 38.2%; Score 47; DB 2; Length 487;
Best Local Similarity 50.0%; Pred. No. 1.2e+02;
Matches 11; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 5 ABGTPLSTFLGLIASRLRSVA 26
DB 394 AVGTATPSYELGLANRLIRVA 415

RESULT 13
Q8WSD1 ID Q8WSD1 PRELIMINARY; PRT; 491 AA.
AC Q8WSD1;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Putative copia-like retrotransposon polyprotein, 5'-partial
DE (Fragment).
GN Name=OSJNB0008A05.28;
OS Oryza sativa (Rice)
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Buehl C.R., Yuan Q., Ouyang S., Liu J., Moffat K.S., Hill J.N.,
RA Gansberger K., Brenner M., Burgess S., Hance M., Shvartsbeyn M.,
RA Tsitrin I., Riggs F., Hsiao J., Zismann V., Blunt S., Pal G.,
RA Vanaken S.E., Utterback T.R., Feldblyum T.V., Kalb E., Quackenbush J.,
RA Salzberg S.L., White O., Fraser C.M.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.

```

```

RN [2]
RP SEQUENCE FROM N.A.
RA Buell R.; (NOV-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; AC091749; AAL31076.1; -.
DR Gramene; Q8W5D1; -.
KW Polyprotein.
FT NON_TPR
SQ SEQUENCE 491 AA; 55838 MW; 3C18C544FCEFE397 CRC64;

Query Match 38.2%; Score 47; DB 2; Length 491;
Best Local Similarity 57.9%; Pred. No. 1.2e+02;
Matches 11; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 6 EGTPPLSTFLGLIASRLRS 24
Db 79 ENVFLSKFHSNAGSLRS 97

RESULT 14
Q38860
ID Q38860 PRELIMINARY; PRT; 506 AA.
AC Q38860
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Fatty acid elongase 1.
GN Name=Pael; Synonyms=AT4g34520, T4L20.100;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eustoids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WS.
RA James D.W. Jr.; Lim E.; Keller J.; Plooy I.; Ralston E.; Dooner H.K.;
RA MEDLINE=9252821; PubMed=7734965;
RT "Directed tagging of the Arabidopsis FATTY ACID ELONGATION1 (FAE1)
RT gene with the maize transposon activator.";
RL Plant Cell 7:309-319(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA Bevan M.; Terry N.; Ardiles W.; Buysshaert C.; Dasseville R.;
RA De Clerck R.; De Keyser A.; Neyt P.; Rouze P.; Van Den Daele H.;
RA Villarroel R.; Gielen J.; Van Montagu M.; Jesse T.; Heijnen L.; Vos P.;
RA Hoheisel J.; Mewes H.W.; Mayer K.F.X.; Lemcke K.; Schueller C.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Terry N.; Ardiles W.; Buysshaert C.; Dasseville R.; De Clerck R.;
RA De Keyser A.; Neyt P.; Rouze P.; Van Den Daele H.; Villarroel R.;
RA Gielen J.; Van Montagu M.; Mewes H.W.; Lemcke K.; Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; U29142; AAA70154.1; -.
DR EMBL; AL023094; CAA18831.1; -.
DR EMBL; AL161595; CAB80169.1; -.
DR PIR; T05272; T05272.
DR GO; GO:0008415; F:acetyltransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0009058; P:biosynthesis; IEA.
DR InterPro; IPR001099; N-C synthase.
DR Pfam; PF02797; Chal_sti_syn; I.
SQ SEQUENCE 506 AA; 56263 MW; 4516D0EF9E8453D18 CRC64;

Query Match 38.2%; Score 47; DB 2; Length 506;
Best Local Similarity 71.4%; Pred. No. 1.2e+02;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 9 FPLSTFLGLIASRL 22
Db 23 FPLTAFLAGKASRL 36

RESULT 15
Q6BTK1
ID Q6BTK1 PRELIMINARY; PRT; 532 AA.
AC Q6BTK1
DT 01-OCT-2004 (TrEMBLrel. 28, Created)
DT 01-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Similar to tr|Q8XLY6 Debaryomyces occidentalis SCRI protein.
GN ORFNames=DEHAOC191189;
OS Debaryomyces hansenii (Yeast) (Torulaspora hansenii).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Debaryomyces.
OX NCBI_TaxID=4959;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CBS767;
RG GENOLEVURES;
RA Dujon B.; Sherman D.; Fischer G.; Durrens P.; Casaregola S.;
RA Lafontaine I.; de Montigny J.; Marck C.; Neuvéglise C.; Talla E.;
RA Coffard N.; Frangeul L.; Aigle M.; Anthouard V.; Babour A.; Barbe V.;
RA Barnay S.; Blanchin S.; Beckerich J.M.; Beyne E.; Bleykasten C.;
RA Boisrame A.; Boyer J.; Cattolico L.; Confanieri F.; de Daruvar A.;
RA Despons L.; Fabre E.; Fairhead C.; Ferry-Dumazet H.; Groppi A.;
RA Hantraye F.; Hennequin C.; Jauniaux N.; Joyet P.; Kachouri R.;
RA Kerrest A.; Koszul R.; Lemaire M.; Lesur I.; Ma L.; Muller H.;
RA Nicaud J.M.; Nikolski M.; Oztas S.; Ozier-Kalogeropoulos O.;
RA Pellenz S.; Potier S.; Richard G.F.; Straub M.L.; Suleau A.;
RA Swennene D.; Tekala F.; Wesolowski-Louvel M.; Westhof E.; Wirth B.;
RA Zenliou-Meyer M.; Zivanovic I.; Bolotin-Fukuhara M.; Thierry A.;
RA Bouchier C.; Caudron B.; Scarpelli C.; Gaillardin C.; Weissenbach J.;
RA Wincker P.; Souciet J.L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CBS767;
RA Genoscope;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR382135; CAG86550.1; -.
SQ SEQUENCE 532 AA; 59070 MW; 47DC9639E0F25FE3 CRC64;

Query Match 38.2%; Score 47; DB 2; Length 532;
Best Local Similarity 52.9%; Pred. No. 1.3e+02;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 GYSEAEGETPPLSTFLIG 17
Db 115 GYSEIVATLPLTLFVIG 131

Search completed: November 10, 2004, 13:38:51
Job time : 34.7478 secs

```

This Page Blank (uspto)

RESULT 2
US-09-583-110-3268
Sequence 3268, Application US/09583110
Patent No. 6699703
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al.
TITLE OF INVENTION: Nucleic Acid and Anti-
Body for Detection and Treatment of
TITLE OF INVENTION: Pneumoniae for Diagn-
osis and Treatment
FILE REFERENCE: PATH00-07A
CURRENT APPLICATION NUMBER: US/09/583,110
CURRENT FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/107,433
PRIOR FILING DATE: 1998-06-30
PRIOR APPLICATION NUMBER: US 60/085,131
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: US 60/051,553
PRIOR FILING DATE: 1997-07-02
NUMBER OF SEQ ID NOS: 5322
SEQ ID NO 3268
LENGTH: 339

; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-3268

Query Match 37.5%; Score 58.5; DB 4; Length 339;
Best Local Similarity 38.7%; Pred. No. 0.66;
Matches 12; Conservative 9; Mismatches 9; Indels 1; Gaps 1;

QY 1 RAPRFKQILLDL-KREIDFNRLVEYFNPL 30
Db 152 RTKLMLEELTDIYKADSEWNVLLRYFNPI 182

RESULT 3

US-09-489-039A-13195
; Sequence 13195, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 13195
; LENGTH: 363
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13195

Query Match 36.2%; Score 56.5; DB 4; Length 363;
Best Local Similarity 42.3%; Pred. No. 1.4;
Matches 11; Conservative 9; Mismatches 5; Indels 1; Gaps 1;

QY 6 IKQILLDL-KREIDFNRLVEYFNPL 30
Db 181 VEQILTDLQAPQEWISALLRYFNVP 206

RESULT 4

US-09-248-796A-17253
; Sequence 17253, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:

; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 17253
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-17253

Query Match 35.3%; Score 55; DB 4; Length 358;
Best Local Similarity 45.8%; Pred. No. 2.4;
Matches 11; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 1 RAPRFKQILLDL-KREIDFNRLV 24
Db 179 RANPISQLFDQKPSDINNV 202

RESULT 5

US-08-861-464-4
; Sequence 4, Application US/08861464
; Patent No. 5874210
; GENERAL INFORMATION:
; APPLICANT: Guarante, Leonard P.
; APPLICANT: Austriaco Jr., Nicanor
; APPLICANT: Kennedy, Brian
; TITLE OF INVENTION: Genes Determining Cellular Senescence
; TITLE OF INVENTION: in yeast
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02173

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/861,464
; FILING DATE: 22-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/396,001
; FILING DATE: 28-FEB-1995
; PRIOR APPLICATION DATA: PCT/US94/09351
; APPLICATION NUMBER: 15-AUG-1994
; FILING DATE: 15-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/107,408
; FILING DATE: 16-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: MIT-6408A2Z
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 781-861-6240
; TELEFAX: 781-861-9540
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 834 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-861-464-4.

Query Match 34.0%; Score 53; DB 2; Length 834;
Best Local Similarity 55.6%; Pred. No. 13;
Matches 15; Conservative 4; Mismatches 6; Indels 2; Gaps 2;

QY 5 FIKQILLDLKREIDFNRLVEYFNPLS 31
Db 392 YIIQFLDIK-ELDFVL-LAEFLNELS 416

RESULT 6

US-08-396-001-4
; Sequence 4, Application US/08396001
; Patent No. 5919618
; GENERAL INFORMATION:

; APPLICANT: Guarante, Leonard P.
; APPLICANT: Austriaco Jr., Nicanor
; APPLICANT: Claus, James
; APPLICANT: Cole, Francesca
; APPLICANT: Kennedy, Brian
; TITLE OF INVENTION: Genes Determining Cellular Senescence in
; TITLE OF INVENTION: Yeast
; NUMBER OF SEQUENCES: 15

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
;; STREET: Two Militia Drive
;; CITY: Lexington
;; STATE: MA
;; COUNTRY: USA
;; ZIP: 02173
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/396,001
;; FILING DATE: 28-FEB-1995
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Granahan, Patricia
;; REGISTRATION NUMBER: 32,227
;; REFERENCE/DOCKET NUMBER: MIT-6408A2
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 617-861-6240
;; TELEFAX: 617-861-9540
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 834 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-396-001-4

Query Match 34.0%; Score 53; DB 2; Length 834;
Best Local Similarity 55.6%; Pred. No. 13;
Matches 15; Conservative 4; Mismatches 6; Indels 2; Gaps 2;

QY 5 FIKOILLDKREIDFNVRVFNPLS 31
DB 392 YIIQFLLDIK-ELDFYL-LAEFNRLS 416

RESULT 7
US-09-323-433A-4
; Sequence 4, Application US/09323433A
; Patent No. 6218512
; GENERAL INFORMATION:
; APPLICANT: Guarente, Leonard P.
; APPLICANT: Austriaco Jr., Nicanor
; APPLICANT: Claus, James J.
; APPLICANT: Cole, Francesca
; APPLICANT: Kennedy, Brian
; TITLE OF INVENTION: GENES DETERMINING CELLULAR SENESENCE IN
; FILE REFERENCE: 0050.1491-003
; CURRENT APPLICATION NUMBER: US/09/323.433A
; CURRENT FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: US 08/396,001
; PRIOR FILING DATE: 1995-02-28
; PRIOR APPLICATION NUMBER: PCT/US94/09351
; PRIOR FILING DATE: 1994-08-15
; PRIOR APPLICATION NUMBER: US 08/107,408
; PRIOR FILING DATE: 1993-08-16
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 834
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-323-433A-4

Query Match 34.0%; Score 53; DB 3; Length 834;
Best Local Similarity 55.6%; Pred. No. 13;
Matches 15; Conservative 4; Mismatches 6; Indels 2; Gaps 2;

QY 5 FIKOILLDKREIDFNVRVFNPLS 31
DB 392 YIIQFLLDIK-ELDFYL-LAEFNRLS 416

RESULT 8
US-09-826-752-4
; Sequence 4, Application US/09826752
; Patent No. 6787300
; GENERAL INFORMATION:
; APPLICANT: Guarente, Leonard P.
; APPLICANT: Austriaco Jr., Nicanor
; APPLICANT: Claus, James J.
; APPLICANT: Cole, Francesca
; APPLICANT: Kennedy, Brian
; TITLE OF INVENTION: GENES DETERMINING CELLULAR SENESENCE IN
; FILE REFERENCE: 0050.1491-005
; CURRENT APPLICATION NUMBER: US/09/826,752
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: US 08/396,001
; PRIOR FILING DATE: 1995-02-28
; PRIOR APPLICATION NUMBER: PCT/US94/09351
; PRIOR FILING DATE: 1994-08-15
; PRIOR APPLICATION NUMBER: US 08/107,408
; PRIOR FILING DATE: 1993-08-16
; PRIOR APPLICATION NUMBER: US 09/323,433
; PRIOR FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 834
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-826-752-4

Query Match 34.0%; Score 53; DB 4; Length 834;
Best Local Similarity 55.6%; Pred. No. 13;
Matches 15; Conservative 4; Mismatches 6; Indels 2; Gaps 2;

QY 5 FIKOILLDKREIDFNVRVFNPLS 31
DB 392 YIIQFLLDIK-ELDFYL-LAEFNRLS 416

RESULT 9
US-09-723-546-9
; Sequence 9, Application US/09723546
; Patent No. 6737259
; GENERAL INFORMATION:
; APPLICANT: Clark, Mike A
; TITLE OF INVENTION: Modified Arginine Deiminase
; FILE REFERENCE: PH00059
; CURRENT APPLICATION NUMBER: US/09/723,546
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/023,809
; PRIOR FILING DATE: 1998-02-13
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 9
; LENGTH: 409
; TYPE: PRT
; ORGANISM: Borrelia afzelii
US-09-723-546-9

Query Match 33.0%; Score 51.5; DB 4; Length 409;
Best Local Similarity 34.5%; Pred. No. 9.6;
Matches 10; Conservative 10; Mismatches 8; Indels 1; Gaps 1;

QY 4 RFKIOILLDKREIDFNVRVFNPLS 31
DB 89 KFISQFILEAEIKDTFTINLLKDFSSLT 117

```

; TITLE OF INVENTION: LEGIONELLA PNEUMOPHILA
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross P.C.
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: U.S.A.
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/766,858A
; APPLICATION NUMBER: US/08/766,858A
; FILING DATE: 13-DEC-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/011,545
; FILING DATE: 13-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Crook, Wanneil W.
; REGISTRATION NUMBER: 31,071
; REFERENCE/DOCKET NUMBER: 3501-4-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 575 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-766-858A-5

Query Match 30.4%; Score 47.5; DB 2; Length 575;
Best Local Similarity 45.8%; Pred. No. 58;
Matches 11; Conservative 6; Mismatches 6; Indels 1; Gaps 1;

QY 8 QILLDLKREIDFNVRVLYVEYFNPLS 31
Db 551 ELLLSLHSHKINRNIKEY-NPLS 573

RESULT 13
US-09-124-141-9
; Sequence 9, Application US/09124141
; Patent No. 6211352
; GENERAL INFORMATION:
; APPLICANT: Harrison, Leonard
; APPLICANT: Honeyman, Margot
; APPLICANT: Cram, David
; APPLICANT: De Aizpua, Henry
; TITLE OF INVENTION: A METHOD FOR THE DIAGNOSIS AND TREATMENT OF GLUTAMIC
; FILE REFERENCE: Phillips Ormonde & Fitzpatrick
; CURRENT APPLICATION NUMBER: US/09/124,141
; CURRENT FILING DATE: 1998-07-29
; EARLIER APPLICATION NUMBER: 08/308,952
; EARLIER FILING DATE: 1994-09-20
; EARLIER APPLICATION NUMBER: 07/839,805
; EARLIER FILING DATE: 1992-02-21
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 9
; LENGTH: 203
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: N-terminal
; OTHER INFORMATION: fragment of mouse brain GAD (MBGAD12)
; US-09-124-141-9

; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1756
; LENGTH: 338
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
; US-09-710-279-1756

Query Match 32.1%; Score 50; DB 4; Length 338;
Best Local Similarity 31.2%; Pred. No. 13;
Matches 10; Conservative 7; Mismatches 7; Indels 8; Gaps 1;

QY 5 FIKQILLDLKRE-----IDFNVRVLYVEYFN 28
Db 39 FVRDLVQDVQRDFMTMAIAMDNLNRYLYDFN 70

RESULT 11
US-09-134-001C-3202
; Sequence 3202, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3202
; LENGTH: 406
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
; US-09-134-001C-3202

Query Match 32.1%; Score 50; DB 3; Length 406;
Best Local Similarity 31.2%; Pred. No. 16;
Matches 10; Conservative 7; Mismatches 7; Indels 7; Gaps 1;

QY 5 FIKQILLDLKRE-----IDFNVRVLYVEYFN 28
Db 107 FVRDLVQDVQRDFMTMAIAMDNLNRYLYDFN 138

RESULT 12
US-08-766-858A-5
; Sequence 5, Application US/08766858A
; Patent No. 5935782
; GENERAL INFORMATION:
; APPLICANT: Cianciotto, Nicholas P.
; APPLICANT: Hickey, Erin K.
; APPLICANT: O'Connell, William A.
; TITLE OF INVENTION: METHODS AND MATERIALS FOR DETECTING
```


Query Match 30.1%; Score 47; DB 3; Length 250;
Best Local Similarity 35.5%; Pred. No. 26;
Matches 11; Conservative 7; Mismatches 11; Indels 2; Gaps 1;

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 16:36:12 ; Search time 33.3022 Seconds
(without alignments)
328.807 Million cell updates/sec

Title: US-10-092-750-43

Perfect score: 156

Sequence: 1 RAPRFKQILLDLKREIDFNVLVEYFNPLS 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	156	100.0	31	14	US-10-092-750-43
2	86	55.1	167	9	US-09-864-761-34496
3	78.5	50.3	1031	14	US-10-082-828A-257
4	78.5	50.3	1275	14	US-10-025-201-3
5	74	47.4	363	14	US-10-012-600B-221
6	73.5	47.1	1274	16	US-10-415-615-3
7	72.5	46.5	1272	15	US-10-114-270-30
8	65.5	42.0	148	14	US-10-104-047-2971
9	61.5	39.4	78	9	US-09-867-550-1462
10	60.5	38.8	338	9	US-09-318-271-8
11	58	37.2	261	15	US-10-424-599-214640
12	57.5	36.9	216	15	US-10-425-114-44206
13	57.5	36.9	351	17	US-10-739-930-8963

14	55	35.3	337	16	US-10-451-467A-588	Sequence 588, App
15	53.5	34.3	351	14	US-10-060-275-2	Sequence 2, Appl1
16	53	34.0	101	15	US-10-424-599-178842	Sequence 178842, A
17	53	34.0	281	15	US-10-282-122A-50612	Sequence 50612, A
18	53	34.0	732	14	US-10-369-493-5349	Sequence 5349, App
19	53	34.0	834	9	US-09-826-752-4	Sequence 4, Appl1
20	52.5	33.7	350	15	US-10-424-599-179426	Sequence 179426, A
21	52.5	33.7	363	15	US-10-425-114-47644	Sequence 47644, A
22	51.5	33.0	207	15	US-10-424-599-232005	Sequence 232005, A
23	51.5	33.0	207	15	US-10-425-114-45766	Sequence 45766, A
24	51.5	33.0	409	16	US-10-674-666-16	Sequence 16, Appl1
25	51.5	33.0	441	15	US-10-424-599-232008	Sequence 232008, A
26	50.5	32.4	348	9	US-09-318-271-10	Sequence 10, Appl1
27	50.5	32.4	348	14	US-10-303-664A-18	Sequence 18, Appl1
28	50	32.1	50	15	US-10-424-599-210612	Sequence 210612, A
29	50	32.1	275	17	US-10-739-930-6227	Sequence 6227, App
30	50	32.1	400	15	US-10-282-122A-70590	Sequence 70590, A
31	49	31.4	171	16	US-10-437-963-105103	Sequence 105103, A
32	49	31.4	751	14	US-10-369-493-9755	Sequence 9755, App
33	48.5	31.1	339	15	US-10-450-677A-1	Sequence 1, Appl1
34	48.5	31.1	348	17	US-10-739-930-5752	Sequence 5752, App
35	48.5	31.1	470	15	US-10-072-012-577	Sequence 577, App
36	48.5	31.1	944	15	US-10-072-012-202	Sequence 202, App
37	48	30.8	85	17	US-10-425-115-213844	Sequence 213844, A
38	48	30.8	169	15	US-10-425-114-46883	Sequence 46883, A
39	48	30.8	212	15	US-10-335-977-6370	Sequence 6370, App
40	48	30.8	267	16	US-10-437-963-159097	Sequence 159097, A
41	48	30.8	278	15	US-10-425-114-46425	Sequence 46425, A
42	48	30.8	286	15	US-10-282-122A-47711	Sequence 47711, A
43	48	30.8	322	16	US-10-437-963-104894	Sequence 104894, A
44	48	30.8	323	14	US-10-369-493-1326	Sequence 1326, App
45	48	30.8	323	14	US-10-369-493-20391	Sequence 20391, A

ALIGNMENTS

RESULT 1

US-10-092-750-43
; Sequence 43, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; APPLICANT: Wright, Martin C.
; TITLE OF INVENTION: Polypeptides Interactive with BCL-X1
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092,750
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/274,526
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-43

Query Match 100.0%; Score 156; DB 14; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.5e-15;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAPRFKQILLDLKREIDFNVLVEYFNPLS 31
DB 1 RAPRFKQILLDLKREIDFNVLVEYFNPLS 31

RESULT 2

US-09-864-761-34496
; Sequence 34496, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:

; APPLICANT: Penn, Sharon G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wensheng
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 ; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
 ; FILE REFERENCE: Aemica-X-1
 ; CURRENT APPLICATION NUMBER: US/09/864,761
 ; CURRENT FILING DATE: 2001-05-23
 ; PRIOR APPLICATION NUMBER: US 60/180,312
 ; PRIOR FILING DATE: 2000-02-04
 ; PRIOR APPLICATION NUMBER: US 60/207,456
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: US 09/632,366
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: GB 24263.6
 ; PRIOR FILING DATE: 2000-10-04
 ; PRIOR APPLICATION NUMBER: US 60/236,359
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00668
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00662
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00661
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00670
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: US 60/234,687
 ; PRIOR FILING DATE: 2000-09-21
 ; PRIOR APPLICATION NUMBER: US 09/608,408
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: US 09/774,203
 ; PRIOR FILING DATE: 2001-01-29
 ; NUMBER OF SEQ ID NOS: 49117
 ; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO 34496
 ; LENGTH: 167
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO AC006459.2
 ; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.5
 ; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.1
 ; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2
 ; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.5
 ; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 6.9
 ; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.6
 ; OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 17
 ; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 6.3
 ; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 11
 ; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.5
 ; OTHER INFORMATION: EST HUMAN HIT: AW562326.1, EVALUATE 4.00e-59
 ; OTHER INFORMATION: SWISSPROT HIT: P08548, EVALUATE 1.00e-51
 ; US-09-864-761-34496
 Query Match 55.1%; Score 86; DB 9; Length 167;
 Best Local Similarity 59.3%; Pred. No. 0.00016;
 Matches 16; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Db 24 APFKIKQLLDLRNEIDSNIVGVNFN 50
 RESULT 3
 US-10-082-828A-257
 ; Sequence 257, Application US/10082828A
 ; Publication No. US20030175715A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Sun, Yongming
 ; APPLICANT: Recipon, Herve
 ; APPLICANT: Salceda, Susana
 ; APPLICANT: Liu, Chenghua
 ; APPLICANT: Turner, Leah
 ; TITLE OF INVENTION: Compositions and Methods Relating to Breast Specific Genes and Prc
 ; FILE REFERENCE: DEX-0247
 ; CURRENT APPLICATION NUMBER: US/10/082,828A
 ; CURRENT FILING DATE: 2002-07-09
 ; PRIOR APPLICATION NUMBER: US 60/243,805
 ; PRIOR FILING DATE: 2000-10-27
 ; NUMBER OF SEQ ID NOS: 266
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 257
 ; LENGTH: 1031
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-082-828A-257
 Query Match 50.3%; Score 78.5; DB 14; Length 1031;
 Best Local Similarity 54.8%; Pred. No. 0.014; 7; Mismatches 6; Indels 1; Gaps 1;
 Matches 17; Conservative
 QY 2 APFKIKQLLDLRNEIDSNIVGVNFN-PLS 31
 Db 23 APFKIKQLLDLRNEIDSNIVGVNFN-PLS 53
 RESULT 4
 US-10-025-201-3
 ; Sequence 3, Application US/10025201
 ; Publication No. US20030003468A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Crow, Mary K.
 ; TITLE OF INVENTION: MARKERS FOR DISEASE SUSCEPTIBILITY AND TARGETS FOR THERAPY
 ; FILE REFERENCE: 5983/ZHS67
 ; CURRENT APPLICATION NUMBER: US/10/025,201
 ; CURRENT FILING DATE: 2001-12-19
 ; PRIOR APPLICATION NUMBER: 60/256,673
 ; PRIOR FILING DATE: 2000-12-19
 ; NUMBER OF SEQ ID NOS: 15
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 3
 ; LENGTH: 1275
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; PUBLICATION INFORMATION:
 ; DATABASE ACCESSION NUMBER: GenBank Accession No. US20030003468A1 U09116
 ; DATABASE ENTRY DATE: 1995-02-02
 ; RELEVANT RESIDUES: (1)..(1275)
 ; US-10-025-201-3
 Query Match 50.3%; Score 78.5; DB 14; Length 1275;
 Best Local Similarity 54.8%; Pred. No. 0.018; 7; Mismatches 6; Indels 1; Gaps 1;
 Matches 17; Conservative
 QY 2 APFKIKQLLDLRNEIDSNIVGVNFN-PLS 31
 Db 121 APFKIKQLLDLRNEIDSNIVGVNFN-PLS 151
 RESULT 5
 US-10-012-600B-221
 ; Sequence 221, Application US/10012600B
 ; Publication No. US20030064377A1

```

; GENERAL INFORMATION:
; APPLICANT: Sun, Yongming
; APPLICANT: Recipon, Hervé
; APPLICANT: Chen, Sei-Yu
; TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific Genes and
; FILE REFERENCE: DEX-0265
; CURRENT APPLICATION NUMBER: US/10/012,600B
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/246,109
; PRIOR FILING DATE: 2000-11-06
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 221
; LENGTH: 363
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-012-600B-221

Query Match 47.4%; Score 74; DB 14; Length 363;
Best Local Similarity 51.7%; Pred. No. 0.021;
Matches 15; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

Qy 2 APRFIQVLDLKRDEIDFNRLVEYFNPL 30
Db 120 APRFIQVLDLKRDEIDFNRLVEYFNPL 148

RESULT 6
US-10-415-615-3
; Sequence 3, Application US/10415615
; Publication No. US20040101943A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: YUE, Henry
; APPLICANT: LU, Yan
; APPLICANT: DING, Li
; APPLICANT: TANG, Y. Tom
; APPLICANT: GANDHI, Ameera R.
; APPLICANT: HAFALIA, April J.A.
; APPLICANT: LAL, Preeti G.
; TITLE OF INVENTION: NUCLEIC ACID MODIFICATION ENZYMES
; FILE REFERENCE: PI-0280 USN
; CURRENT APPLICATION NUMBER: US/10/415,615
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: PCT/US01/46301
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: US 60/245,458
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: US 60/255,107
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PERL Program
; SEQ ID NO 3
; LENGTH: 1274
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 1258887CD1
; US-10-415-615-3

Query Match 47.1%; Score 73.5; DB 16; Length 1274;
Best Local Similarity 53.3%; Pred. No. 0.095;
Matches 16; Conservative 7; Mismatches 6; Indels 1; Gaps 1;

Qy 3 PRFIQVLDLKRDEIDFNRLVEYFN-ELS 31
Db 122 PRFIQVLDLKRDEIDFNRLVEYFN-ELS 151

RESULT 7
US-10-114-270-30

```

```

; Sequence 30, Application US/10114270
; Publication No. US20040030110A1
; GENERAL INFORMATION:
; APPLICANT: Guo, Xiaojia
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Miller, Charles E.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Patturajan, Meera
; APPLICANT: Liu, Ziaohong
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Li, Li
; APPLICANT: Vernet, Corine
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Gorman, Linda
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Smithson, Glennda
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Shinkets, Richard A.
; APPLICANT: Gangolli, Bsha A.
; APPLICANT: Taubier Jr., Raymond J.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Ji, Weizhen
; APPLICANT: Anderson, David W.
; APPLICANT: Liete, Mario W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Stone, David J.
; APPLICANT: MacDougall, John R.
; APPLICANT: Rothenberg, Mark E.
; TITLE OF INVENTION: No. US20040030110A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-322C
; CURRENT APPLICATION NUMBER: US/10/114,270
; CURRENT FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/281,086
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/281,136
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/281,863
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/281,906
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/282,020
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,930
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 60/282,934
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 60/283,512
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/283,710
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,234
; PRIOR FILING DATE: 2001-04-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 470
; SEQ ID NO 30
; LENGTH: 1272
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-114-270-30

Query Match 46.5%; Score 72.5; DB 15; Length 1272;
Best Local Similarity 51.8%; Pred. No. 0.13;
Matches 16; Conservative 7; Mismatches 7; Indels 1; Gaps 1;

Qy 2 APRFIQVLDLKRDEIDFNRLVEYFN-ELS 31
Db 117 APRFIQVLDLKRDEIDFNRLVEYFN-ELS 147

```

RESULT 8
 US-10-104-047-2971
 ; Sequence 2971, Application US/10104047
 ; Publication No. US20030236392A1
 ; GENERAL INFORMATION:
 ; APPLICANT: HELIX RESEARCH INSTITUTE
 ; TITLE OF INVENTION: NO. US20030236392A1el full length cDNA
 ; FILE REFERENCE: H1-A0105
 ; CURRENT APPLICATION NUMBER: US/10/104,047
 ; CURRENT FILING DATE: 2002-03-25
 ; PRIOR APPLICATION NUMBER:
 ; PRIOR FILING DATE:
 ; NUMBER OF SEQ ID NOS: 4096
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2971
 ; LENGTH: 148
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-104-047-2971

Query Match 42.0%; Score 65.5; DB 14; Length 148;
 Best Local Similarity 51.6%; Pred. No. 0.14;
 Matches 16; Conservative 5; Mismatches 9; Indels 1; Gaps 1;
 QY 2 APRFKQILLDLKREIDFNRLVEYFN-PLS 31
 Db 103 APRLIKQVLRDRQRDLDSHTITVGFNPLS 133

RESULT 9
 US-09-867-550-1462
 ; Sequence 1462, Application US/09867550
 ; Patent No. US20020082206A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Leach, Martin D.
 ; APPLICANT: Mehraban, Fuad,
 ; APPLICANT: Conley, Pamela
 ; APPLICANT: Law, Debbie
 ; APPLICANT: Topper, James
 ; TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells and
 ; TITLE OF INVENTION: Theraaby
 ; FILE REFERENCE: 21402-013 (Cura-313)
 ; CURRENT APPLICATION NUMBER: US/09/867,550
 ; CURRENT FILING DATE: 2001-03-20
 ; PRIOR APPLICATION NUMBER: US9N 60/208,427
 ; PRIOR FILING DATE: 2000-05-30
 ; NUMBER OF SEQ ID NOS: 2125
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1462
 ; LENGTH: 78
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; NAME/KEY: VARIANT
 ; LOCATION: (1)
 ; OTHER INFORMATION: Wherein Xaa may be any one of Ala or Arg or Asn or Asp or Cys or
 ; OTHER INFORMATION: His or Ile or Leu or Phe or Pro or Ser or Thr or Tyr or Val
 US-09-867-550-1462

Query Match 39.4%; Score 61.5; DB 9; Length 78;
 Best Local Similarity 57.1%; Pred. No. 0.27;
 Matches 16; Conservative 2; Mismatches 9; Indels 1; Gaps 1;
 QY 1 RAPRFKQILLDLKREIDFNRLVEYFN 28
 Db 19 RAPRYIKQI-LDLRRVIDSKTKADEFN 45

RESULT 10
 US-09-318-271-8
 ; Sequence 8, Application US/09318271A
 ; Patent No. US20020012979A1

; GENERAL INFORMATION:
 ; APPLICANT: Berry, Alan
 ; APPLICANT: Running, Jeffrey A.
 ; APPLICANT: Severson, David K.
 ; APPLICANT: Burlingame, Richard P.
 ; TITLE OF INVENTION: "VITAMIN C PRODUCTION IN MICROORGANISMS AND PLANTS"
 ; FILE REFERENCE: 3161-24
 ; CURRENT APPLICATION NUMBER: US/09/318,271A
 ; CURRENT FILING DATE: 1999-05-25
 ; EARLIER APPLICATION NUMBER: 60/125,073
 ; EARLIER FILING DATE: 1999-03-17
 ; EARLIER APPLICATION NUMBER: 60/125,054
 ; EARLIER FILING DATE: 1999-03-18
 ; EARLIER APPLICATION NUMBER: 60/088,549
 ; EARLIER FILING DATE: 1998-06-08
 ; NUMBER OF SEQ ID NOS: 15
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 8
 ; LENGTH: 338
 ; TYPE: PRT
 ; ORGANISM: Escherichia coli
 US-09-318-271-8

Query Match 38.8%; Score 60.5; DB 9; Length 338;
 Best Local Similarity 46.2%; Pred. No. 1.8;
 Matches 12; Conservative 8; Mismatches 5; Indels 1; Gaps 1;
 QY 6 IKQILLDL-KREIDFNRLVEYFNPL 30
 Db 156 VEQILTDLQAPDWSIALRYFN 181

RESULT 11
 US-10-424-599-214640
 ; Sequence 214640, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Kovalic David K
 ; APPLICANT: Zhou Yihua
 ; APPLICANT: Cao Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 214640
 ; LENGTH: 261
 ; TYPE: PRT
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT3847_35847C.1.pap
 US-10-424-599-214640

Query Match 37.2%; Score 58; DB 15; Length 261;
 Best Local Similarity 35.7%; Pred. No. 3.2;
 Matches 10; Conservative 8; Mismatches 10; Indels 0; Gaps 0;
 QY 1 RAPRFKQILLDLKREIDFNRLVEYFN 28
 Db 222 KAKKWAEBIILDLKSNDSMAVMILRYFN 249

RESULT 12
 US-10-425-114-44206
 ; Sequence 44206, Application US/10425114
 ; Publication No. US20040034888A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E

; PRIOR APPLICATION NUMBER: EP 01870003.9

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 12:29:32 ; Search time 6.90674 Seconds
(without alignments)
431.857 Million cell updates/sec

Title: US-10-092-750-43
Perfect score: 156
Sequence: 1 RARFVKQILLDKREIDFNRLVEYFNPLS 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR.79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	78.5	50.3	1275	2 B28096	line-1 protein ORF
2	78.5	50.3	1275	2 138558	reverse transcript
3	78.5	50.3	1275	2 S65824	reverse transcript
4	77.5	49.7	1259	4 GNHUL1	retrovirus-related
5	77.5	49.7	1280	2 B34087	hypothetical prote
6	60.5	38.8	338	1 XOEUCG	UDP-glucose 4-epim
7	60.5	38.8	338	2 C90727	UDP-galactose-4-ep
8	60.5	38.8	338	2 D85578	UDP-galactose-4-ep
9	58.5	37.5	339	2 B95187	UDP-glucose 4-epim
10	58.5	37.5	339	2 C98054	UDP-glucose 4-epim
11	57.5	36.9	354	2 T10436	UDP-glucose 4-epim
12	57	36.5	285	2 S49879	hypothetical narbo
13	57	36.5	285	2 T12157	nodulin - fava bea
14	57	36.5	285	2 S49898	hypothetical narbo
15	56.5	36.2	337	1 A37760	UDP-glucose 4-epim
16	56.5	36.2	338	1 S51328	UDP-glucose 4-epim
17	56.5	36.2	337	2 AF0594	UDP-glucose 4-epim
18	56.5	36.2	347	2 T19989	hypothetical prote
19	56	35.9	285	2 T12156	nodulin, isoform N
20	55.5	35.6	500	2 S16788	probable reverse t
21	55	35.3	796	2 T43782	hypothetical prote
22	54	34.6	431	2 D84779	hypothetical prote
23	53.5	34.3	336	1 S70744	UDP-glucose 4-epim
24	53	34.0	732	2 T32023	hypothetical prote
25	53	34.0	859	2 S64195	HTB1 protein - yea
26	52.5	33.7	350	2 T06526	UDP-glucose 4-epim
27	52.5	33.7	351	2 B86261	UDP-glucose 4-epim
28	52.5	33.7	351	2 S62783	UDP-glucose 4-epim
29	52	33.3	205	2 A72479	hypothetical prote

RESULT 1

B28096
line-1 protein ORF2 - human
C:Species: Homo sapiens (man)
C:Date: 03-Nov-1988 #sequence_revision 03-Nov-1988 #text_change 09-Jul-2004
C:Accession: B28096
R:Skowronski, J.; Fanning, T.G.; Singer, M.F.
Mol. Cell. Biol. 8, 1385-1397, 1988
A:Title: Unit-length line-1 transcripts in human teratocarcinoma cells.
A:Reference number: A28096; MUID:88246405; PMID:2454389
A:Accession: B28096
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-1275 <SKO>
A:Cross-references: UNIPROT:Q15604; UNIPROT:Q9UN80; UNIPROT:Q12881; UNIPROT:O00363; UNIPROT:Q9Y5K0; UNIPROT:O00366; UNIPROT:Q8TE30; UNIPROT:O00375
C:Superfamily: pol polyprotein

Query Match 50.3%; Score 78.5; DB 2; Length 1275;
Best Local Similarity 54.8%; Pred. No. 0.012;
Matches 17; Conservative 7; Mismatches 6; Indels 1; Gaps 1;

QY 2 APRFVKQILLDKREIDFNRLVEYFN-PLS 31
DB 121 APRFVKQVLSLDQLDSDSHLTIMGDFNTPLS 151

RESULT 2

I38588
reverse transcriptase homolog - human retrotransposon L1
N:Alternate names: ORF2 protein
C:Species: Homo sapiens (man)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C:Accession: I38588
R:Holmes, S.E.; Dombroski, B.A.; Krebs, C.M.; Boehm, C.D.; Kazazian, H.H.
Nature Genet. 7, 143-148, 1994
A:Title: A new retrotransposable human L1 element from the LRE2 locus on chromosome 1q p1
A:Reference number: I38587; MUID:95004577; PMID:7920631
A:Accession: I38588
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1275 <RGS>
A:Cross-references: UNIPROT:Q12881; EMBL:U09116; NID:G483914; PIDN:AAB60345.1; PID:G48391
C:Superfamily: pol polyprotein

Query Match 50.3%; Score 78.5; DB 2; Length 1275;
Best Local Similarity 54.8%; Pred. No. 0.012;
Matches 17; Conservative 7; Mismatches 6; Indels 1; Gaps 1;

QY 2 APRFVKQILLDKREIDFNRLVEYFN-PLS 31
DB 121 APRFVKQVLSLDQLDSDSHLTIMGDFNTPLS 151

T518.7 protein - A
UDFglucose 4-epime
UDFglucose 4-epime
transcription-repa
UDFglucose 4-epime
hypothetical prote
hypothetical narbo
UDP-glucose 4-epim
formate hydrogenly
unknown protein T1
conserved hypothat
UDFglucose 4-epime
conserved hypothat
hypothetical prote
UDFglucose 4-epime
probable 3-isoprop

```

RESULT 3
S65824
reverse transcriptase homolog - human transposon L1.1
C:Species: Homo sapiens (man)
C:Date: 20-Jul-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C:Accession: S65824
R:Dombroski, B.A.
submitted to the EMBL Data Library, January 1992
A:Description: Isolation of an active human transposable element.
A:Reference number: S65823
A:Accession: S65824
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1275 <DOM>
A:Cross-references: UNIPROT:Q15604; EMBL:M80340; NID:g339767; PIDN:AAA51622.1; PID:g3397
C:Superfamily: pol polyprotein

Query Match          50.3%; Score 78.5; DB 2; Length 1275;
Best Local Similarity 54.8%; Pred. No. 0.012; 6; Indels 1; Gaps 1;
Matches 17; Conservative 7; Mismatches 7;

QY 2 APRFIKQILLDKREIDFNRLVEYFN-PLS 31
|||||:|||||:|||||:|||||:|||||
Db 121 APRFIKQVLSLDQLDLSHTLIMGDNFTPLS 151

RESULT 4
GNHUL1
retrovirus-related reverse transcriptase pseudogene - human
C:Species: Homo sapiens (man)
C:Date: 31-Mar-1988 #sequence_revision 04-Jan-1996 #text_change 09-Jul-2004
C:Accession: A25313
A:Reference number: A93381; MUID:86230917; PMID:2423883
A:Accession: A25313
A:Status: conceptual translation of pseudogene
A:Molecule type: DNA
A:Residues: 1-1259 <HAT>
A:Cross-references: UNIPROT:P08547
A:Note: this sequence was constructed from an alignment of published and unpublished seq
C:Keywords: reverse transcriptase; pseudogene

Query Match          49.7%; Score 77.5; DB 4; Length 1259;
Best Local Similarity 54.8%; Pred. No. 0.016; 7; Indels 1; Gaps 1;
Matches 17; Conservative 7; Mismatches 7;

QY 2 APRFIKQILLDKREIDFNRLVEYFN-PLS 31
|||||:|||||:|||||:|||||:|||||
Db 120 APRFIKQVLSLDQLDLSHTLIMGDNFTPLS 150

RESULT 5
B34087
hypothetical protein (L1H 3' region) - human
C:Species: Homo sapiens (man)
C:Date: 30-Mar-1990 #sequence_revision 30-Mar-1990 #text_change 09-Jul-2004
C:Accession: B34087
R:Scott, A.F.; Schmeckpeper, B.J.; Abdelrazik, M.; Comey, C.T.; O'Hara, B.; Rossiter, J.
Genomics 1, 113-125, 1987
A:Title: Origin of the human L1 elements: proposed progenitor genes deduced from a conse
A:Reference number: A34087; MUID:86085185; PMID:3652483
A:Accession: B34087
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-1280 <SCO>
A:Cross-references: UNIPROT:Q9Y5K0
C:Superfamily: pol polyprotein

Query Match          49.7%; Score 77.5; DB 2; Length 1280;
Best Local Similarity 54.8%; Pred. No. 0.017; 6; Indels 1; Gaps 1;
Matches 17; Conservative 7; Mismatches 7;

QY 2 APRFIKQILLDKREIDFNRLVEYFN-PLS 31
|||||:|||||:|||||:|||||:|||||
Db 126 APRFIKQVLSLDQLDLSHTLIMGDNFTPLS 156

RESULT 6
XUECUG
UDPglucose 4-epimerase (EC 5.1.3.2) - Escherichia coli (strain K-12)
N:Alternate names: UDPglucose 4-epimerase
C:Species: Escherichia coli
C:Date: 31-Dec-1990 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004
C:Accession: S02089; S00759; S14391; G64811; B25764
R:Lemaire, H.G.
submitted to the EMBL Data Library, April 1988
A:Reference number: S02089
A:Accession: S02089
A:Molecule type: DNA
A:Residues: 1-338 <LEM1>
A:Cross-references: UNIPROT:P09147; EMBL:X06226; NID:g41522; PIDN:CAA29573.1; PID:g41523
R:Lemaire, H.G.; Mueller-Hill, B.
Nucleic Acids Res. 14, 7705-7711, 1986
A:Title: Nucleotide sequences of the gale gene and the galt gene of E. coli.
A:Reference number: S00722; MUID:87040735; PMID:3022232
A:Accession: S00759
A:Molecule type: DNA
A:Residues: 1-294, 'AFRPTGTRAKPTVN', 310-338 <LEM2>
A:Cross-references: EMBL:X06226
A:Note: this sequence has been revised in reference S02089
R:Bernardi, F.; Bernardi, A.
DNA Seq. 1, 147-150, 1990
A:Reference number: S14391; MUID:92190543; PMID:2134186
A:Accession: S14391
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-139, 'LLPIPGQ' <BER>
A:Cross-references: EMBL:X51449; NID:g42412; PIDN:CAA35813.1; PID:g42413
R:Blattner, P.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Col
A:Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A54720; MUID:97426617; PMID:9278503
A:Accession: G64811
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-338 <BLAT>
A:Cross-references: GB:AE000178; GB:U00096; NID:g1786967; PIDN:AAC73846.1; PID:g1786974;
C:Genetics:
A:Gene: gale
A:Map position: 17 min
C:Function:
A:Description: isomerase
A:Pathway: galactose metabolism
C:Superfamily: Escherichia coli UDPglucose 4-epimerase; UDPglucose 4-epimerase homology
C:Keywords: galactose metabolism; isomerase
F:3-336/Domain: UDPglucose 4-epimerase homology <UDP>

Query Match          38.8%; Score 60.5; DB 1; Length 338;
Best Local Similarity 46.2%; Pred. No. 0.95; 8; Mismatches 5; Indels 1; Gaps 1;
Matches 12; Conservative 8;

QY 6 IKQILLDL-KREIDFNRLVEYFNPL 30
|||||:|||||:|||||:|||||:|||||
Db 156 VEQILTDLQKAPDWSIALRYFNVP 181

RESULT 7
C90727
UDP-galactose-4-epimerase [imported] - Escherichia coli (strain O157:H7, substrain RMD (
```

```

Best Local Similarity 54.8%; Pred. No. 0.017;
Matches 17; Conservative 7; Mismatches 6; Indels 1; Gaps 1;

QY 2 APRFIKQILLDKREIDFNRLVEYFN-PLS 31
|||||:|||||:|||||:|||||:|||||
Db 126 APRFIKQVLSLDQLDLSHTLIMGDNFTPLS 156

RESULT 6
XUECUG
UDPglucose 4-epimerase (EC 5.1.3.2) - Escherichia coli (strain K-12)
N:Alternate names: UDPglucose 4-epimerase
C:Species: Escherichia coli
C:Date: 31-Dec-1990 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004
C:Accession: S02089; S00759; S14391; G64811; B25764
R:Lemaire, H.G.
submitted to the EMBL Data Library, April 1988
A:Reference number: S02089
A:Accession: S02089
A:Molecule type: DNA
A:Residues: 1-338 <LEM1>
A:Cross-references: UNIPROT:P09147; EMBL:X06226; NID:g41522; PIDN:CAA29573.1; PID:g41523
R:Lemaire, H.G.; Mueller-Hill, B.
Nucleic Acids Res. 14, 7705-7711, 1986
A:Title: Nucleotide sequences of the gale gene and the galt gene of E. coli.
A:Reference number: S00722; MUID:87040735; PMID:3022232
A:Accession: S00759
A:Molecule type: DNA
A:Residues: 1-294, 'AFRPTGTRAKPTVN', 310-338 <LEM2>
A:Cross-references: EMBL:X06226
A:Note: this sequence has been revised in reference S02089
R:Bernardi, F.; Bernardi, A.
DNA Seq. 1, 147-150, 1990
A:Reference number: S14391; MUID:92190543; PMID:2134186
A:Accession: S14391
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-139, 'LLPIPGQ' <BER>
A:Cross-references: EMBL:X51449; NID:g42412; PIDN:CAA35813.1; PID:g42413
R:Blattner, P.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Col
A:Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A54720; MUID:97426617; PMID:9278503
A:Accession: G64811
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-338 <BLAT>
A:Cross-references: GB:AE000178; GB:U00096; NID:g1786967; PIDN:AAC73846.1; PID:g1786974;
C:Genetics:
A:Gene: gale
A:Map position: 17 min
C:Function:
A:Description: isomerase
A:Pathway: galactose metabolism
C:Superfamily: Escherichia coli UDPglucose 4-epimerase; UDPglucose 4-epimerase homology
C:Keywords: galactose metabolism; isomerase
F:3-336/Domain: UDPglucose 4-epimerase homology <UDP>

Query Match          38.8%; Score 60.5; DB 1; Length 338;
Best Local Similarity 46.2%; Pred. No. 0.95; 8; Mismatches 5; Indels 1; Gaps 1;
Matches 12; Conservative 8;

QY 6 IKQILLDL-KREIDFNRLVEYFNPL 30
|||||:|||||:|||||:|||||:|||||
Db 156 VEQILTDLQKAPDWSIALRYFNVP 181

RESULT 7
C90727
UDP-galactose-4-epimerase [imported] - Escherichia coli (strain O157:H7, substrain RMD (
```

C:\Species: Escherichia coli
C:\Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:\Accession: C90727
C:\Hayashi, T.; Makino, K.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:\Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
A:\Reference number: A96229; PMID:21156231; PMID:11258796
A:\Status: preliminary
A:\Molecule type: DNA
A:\Residues: 1-338 <HAY>
A:\Cross-references: UNIPROT:Q8X942; UNIPROT:Q8FJR9; GB:BA000007; PIDN:BAB34210.1; PID:g1
A:\Experimental source: strain O157:H7, substrain RIMD 050952
C:\Genetics:
A:\Gene: EC50787
C:\Superfamily: Escherichia coli UDPglucose 4-epimerase; UDPglucose 4-epimerase homology

Query Match 38.8%; Score 60.5; DB 2; Length 338;
Best Local Similarity 46.2%; Pred. No. 0.95;
Matches 12; Conservative 8; Mismatches 5; Indels 1; Gaps 1;

Qy 6 IKQILLDL-KREIDFNRLVEYNPL 30
:||||| :||| :||| :|||
Db 156 VEQILTDLQAQPDSIALRYENPV 181

RESULT 8
D85578
UDP-galactose-4-epimerase [imported] - Escherichia coli (strain O157:H7, substrain EDL933)
C:\Species: Escherichia coli
C:\Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:\Accession: D85578
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:\Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:\Reference number: A85480; PMID:21074935; PMID:11206551
A:\Status: preliminary
A:\Molecule type: DNA
A:\Residues: 1-338 <STO>
A:\Cross-references: UNIPROT:Q8X942; UNIPROT:Q8FJR9; GB:AB005174; NID:gl2513688; PIDN:AAG
A:\Experimental source: strain O157:H7, substrain EDL933
C:\Genetics:
A:\Gene: galE
C:\Superfamily: Escherichia coli UDPglucose 4-epimerase; UDPglucose 4-epimerase homology

Query Match 38.8%; Score 60.5; DB 2; Length 338;
Best Local Similarity 46.2%; Pred. No. 0.95;
Matches 12; Conservative 8; Mismatches 5; Indels 1; Gaps 1;

Qy 6 IKQILLDL-KREIDFNRLVEYNPL 30
:||||| :||| :||| :|||
Db 156 VEQILTDLQAQPDSIALRYENPV 181

RESULT 9
B95187
UDP-glucose 4-epimerase [imported] - Streptococcus pneumoniae (strain TIGR4)
C:\Species: Streptococcus pneumoniae
C:\Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
C:\Accession: B95187
R:Tetzelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfel,
nson, T.; Hickley, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A:\Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A:\Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A:\Reference number: A95000; PMID:21357209; PMID:11463916
A:\Status: preliminary
A:\Molecule type: DNA

A;Residues: 1-339 <KUR>
A;Cross-references: UNIPROT:Q97PK2; GB:AE005672; PIDN:AAK75691.1; PID:g14973098; GSPDB:B.6
A;Experimental source: strain TIGR4
C;Genetics:
A;Gene: SPI607
C;Superfamily: Escherichia coli UDPglucose 4-epimerase; UDPglucose 4-epimerase homology

Query Match 37.5%; Score 58.5; DB 2; Length 339;
Best Local Similarity 38.7%; Pred. No. 1.8;
Matches 12; Conservative 9; Mismatches 9; Indels 1; Gaps 1;

Oy 1 RAPRFIKOILLDL-KREIDFNRLVEYFNPL 30
Db 152 RTKLMBEILTDIYKADSEWNVLLRYFNPI 182
| : : : | : : : | : : : | : : : |
| : : : | : : : | : : : | : : : |

RESULT 10
C98054
UDPglucose 4-epimerase (EC 5.1.1.3.2) [imported] - Streptococcus pneumoniae (strain R6)
C;Species: Streptococcus pneumoniae
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C;Accession: C98054
E;Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszczyk, L.; Burgett, S.; DeHoff, B.S.; Eick,
R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; MC
Y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A;Reference number: A97872; MOID:21429245; PMID:11544234
A;Accession: C98054
A;Status: Preliminary
A;Molecule type: DNA
A;Residues: 1-339 <KUR>
A;Cross-references: UNIPROT:Q8DNV6; GB:AE007317; PIDN:AAL00264.1; PID:g15459117; GSPDB:B.6
C;Genetics:
A;Gene: galE
C;Superfamily: Escherichia coli UDPglucose 4-epimerase; UDPglucose 4-epimerase homology
C;Keywords: isomerase

Query Match 37.5%; Score 58.5; DB 2; Length 339;
Best Local Similarity 38.7%; Pred. No. 1.8;
Matches 12; Conservative 9; Mismatches 9; Indels 1; Gaps 1;

Oy 1 RAPRFIKOILLDL-KREIDFNRLVEYFNPL 30
Db 152 RTKLMBEILTDIYKADSEWNVLLRYFNPI 182
| : : : | : : : | : : : | : : : |
| : : : | : : : | : : : | : : : |

RESULT 11
TI0496
UDPglucose 4-epimerase (EC 5.1.1.3.2) (clone GEPI42) - guar
N;Alternate names: UDP-galactose 4-epimerase
C;Species: Cyanopsis tetragonoloba (Guar, cluster bean)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C;Accession: TI0496
R;Brusteadt, J.; Joersbo, M.; Pedersen, S.G.; Marcussen, J.
Plant Sci. 142, 147-154, 1999
A;Title: Isolation and expression of two cDNA clones encoding UDP-galactose epimerase xfr
A;Reference number: Z17058
A;Accession: TI0496
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-354 <BRU>
A;Cross-references: UNIPROT:O65780; EMBL:AJ005091
C;Superfamily: Escherichia coli UDPglucose 4-epimerase; UDPglucose 4-epimerase homology
C;Keywords: Galactose metabolism; isomerase
P;12-350/Domain: UDPglucose 4-epimerase homology <UDP>

Query Match 36.9%; Score 57.5; DB 2; Length 354;
Best Local Similarity 32.3%; Pred. No. 2.6;
Matches 10; Conservative 12; Mismatches 8; Indels 1; Gaps 1;

Ov 1 RAPRFIKOILLDKR-EIDFNRLVEYFNPL 30

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 12:27:34 ; Search time 37.8532 Seconds
(without alignments)
471.205 Million cell updates/sec

Title: US-10-092-750-43

Perfect score: 156

Sequence: 1 RARFKQILLDKREIDFNRLVEYFNPLS 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	78.5	50.3	1192	2	Q7KZ41
2	78.5	50.3	1275	2	O00360
3	78.5	50.3	1275	2	O00362
4	78.5	50.3	1275	2	O00363
5	78.5	50.3	1275	2	O00366
6	78.5	50.3	1275	2	O00368
7	78.5	50.3	1275	2	O00370
8	78.5	50.3	1275	2	O00375
9	78.5	50.3	1275	2	O00378
10	78.5	50.3	1275	2	Q8TE30
11	78	50.0	244	2	Q6ZN38
12	78	50.0	244	2	BAD18477
13	77.5	49.7	1259	1	L1N1 HUMAN
14	77.5	48.4	314	2	O6ZNC3
15	75.5	48.4	314	2	BAD18452
16	74.5	47.8	152	2	Q9Y443
17	72.5	46.5	1275	2	O00372
18	71.5	45.8	202	2	Q6ZND9
19	71.5	45.8	202	2	BAD18436
20	70.5	45.2	177	2	Q6ZNE0
21	70.5	45.2	177	2	BAD18435
22	65.5	42.0	148	2	Q8NAQ0
23	60.5	38.8	264	2	Q6ZNC2
24	60.5	38.8	264	2	BAD18453
25	60.5	38.8	338	1	GALE_ECOLI
26	60.5	38.8	338	2	Q7AG12
27	60.5	38.8	338	2	Q7UDF1
28	60.5	38.8	338	2	Q8FJR9
29	60.5	38.8	338	2	Q8X942
30	60.5	38.8	342	2	Q83SA4
31	59	37.8	338	2	Q87GR1

ALIGNMENTS

RESULT 1

Q7KZ41 PRELIMINARY; PRT; 1192 AA.

AC Q7KZ41; 625 37.8 59 32
 DT 05-JUL-2004 (TREMELrel. 27, Created)
 DT 05-JUL-2004 (TREMELrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMELrel. 27, Last annotation update)
 DE Hypothetical protein (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89233117; PubMed=2497061;
 RA Woods-Samuels P., Wong C., Mathias S.L., Scott A.F.,
 RA Kazanian H.H. Jr., Antonarakis S.E.,
 RT "Characterization of a nondeleterious L1 insertion in an intron of the
 RT human factor VIII gene and further evidence of open reading frames in
 RT functional L1 elements."
 RL Genomics 4:290-296(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Woods-Samuels P.,
 DR EMBL; M22333; AAA88037.1; -
 DR InterPro; IPR005135; Exo_endo_phos.
 DR InterPro; IPR000477; RVTse.
 DR Pfam; PF03372; Exo_endo_phos; 1.
 DR Pfam; PF00078; RVT; 1.
 KW Hypothetical protein; RNA-directed DNA polymerase; Transferase.
 FT NON_TER 1
 SQ SEQUENCE 1192 AA; 139677 MW; D706D841DA50DAD4 CRC64;

Query Match 50.3%; Score 78.5; DB 2; Length 1192;
 Best Local Similarity 54.8%; Pred. No. 0.03;
 Matches 17; Conservative 7; Mismatches 6; Indels 1; Gaps 1;

QY 2 APRFKQILLDKREIDFNRLVEYFN-PLS 31

DB 38 APRFKQVLSLDQLRDLSDHTLMDGDFNPLS 68

RESULT 2

O00360 PRELIMINARY; PRT; 1275 AA.

AC O00360;
 DT 01-JUL-1997 (TREMELrel. 04, Created)
 DT 01-JUL-1997 (TREMELrel. 04, Last sequence update)
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
 DE Putative p150 (Human).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97285120; PubMed=9140393;
RA Sasaman D.M., Dombroski B.A., Moran J.V., Kimberland M.L., Naas T.P.,
RA DeBerardinis R.J., Gabriel A., Swergold G.D., Kazazian H.H. Jr.;
RT "Many human Li elements are capable of retrotransposition.";
RL Nat. Genet. 16:37-43(1997).
DR EMBL; U93563; AAC51261.1; -.
DR PIR; B28096; B28096.
DR PIR; S23650; S23650.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006278; F:RNA-dependent DNA replication; IEA.
DR InterPro; IPR005135; Exo_endo_phos.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF03372; Exo_endo_phos; 1.
DR Pfam; PF00078; RVT; 1.
KW RNA-directed DNA polymerase; Transferase.
SQ SEQUENCE 1275 AA; 149062 MW; 350B4F0358E525F0 CRC64;

Query Match 50.3%; Score 78.5; DB 2; Length 1275;
Best Local Similarity 54.8%; Pred. No. 0.032;
Matches 17; Conservative 7; Mismatches 6; Indels 1; Gaps 1;

QY 2 APRFKQILLDKREIDFNVLVEYFN-PLS 31
DB 121 APRFKQVLSDLQRLDSHTLIMGDFNTPLS 151

RESULT 3
O00362 PRELIMINARY; PRT; 1275 AA.
AC O00362;
DT 01-JUL-1997 (TRENBLrel. 04, Created)
DT 01-JUL-1997 (TRENBLrel. 04, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Putative p150.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97285120; PubMed=9140393;
RA Sasaman D.M., Dombroski B.A., Moran J.V., Kimberland M.L., Naas T.P.,
RA DeBerardinis R.J., Gabriel A., Swergold G.D., Kazazian H.H. Jr.;
RT "Many human Li elements are capable of retrotransposition.";
RL Nat. Genet. 16:37-43(1997).
DR EMBL; U93564; AAC51263.1; -.
DR PIR; B28096; B28096.
DR PIR; S23650; S23650.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006278; F:RNA-dependent DNA replication; IEA.
DR InterPro; IPR005135; Exo_endo_phos.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF03372; Exo_endo_phos; 1.
DR Pfam; PF00078; RVT; 1.
KW RNA-directed DNA polymerase; Transferase.
SQ SEQUENCE 1275 AA; 149201 MW; 23D516D6E4358F28 CRC64;

Query Match 50.3%; Score 78.5; DB 2; Length 1275;
Best Local Similarity 54.8%; Pred. No. 0.032;
Matches 17; Conservative 7; Mismatches 6; Indels 1; Gaps 1;

QY 2 APRFKQILLDKREIDFNVLVEYFN-PLS 31
DB 121 APRFKQVLSDLQRLDSHTLIMGDFNTPLS 151

RESULT 4
O00363 PRELIMINARY; PRT; 1275 AA.
AC O00363;
DT 01-JUL-1997 (TRENBLrel. 04, Created)
DT 01-JUL-1997 (TRENBLrel. 04, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Putative p150.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97285120; PubMed=9140393;
RA Sasaman D.M., Dombroski B.A., Moran J.V., Kimberland M.L., Naas T.P.,
RA DeBerardinis R.J., Gabriel A., Swergold G.D., Kazazian H.H. Jr.;
RT "Many human Li elements are capable of retrotransposition.";
RL Nat. Genet. 16:37-43(1997).
DR EMBL; U93565; AAC51264.1; -.
DR PIR; B28096; B28096.
DR PIR; S23650; S23650.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006278; F:RNA-dependent DNA replication; IEA.
DR InterPro; IPR005135; Exo_endo_phos.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF03372; Exo_endo_phos; 1.
DR Pfam; PF00078; RVT; 1.
KW RNA-directed DNA polymerase; Transferase.
SQ SEQUENCE 1275 AA; 149007 MW; A868978EA3FD8F74 CRC64;

Query Match 50.3%; Score 78.5; DB 2; Length 1275;
Best Local Similarity 54.8%; Pred. No. 0.032;
Matches 17; Conservative 7; Mismatches 6; Indels 1; Gaps 1;

QY 2 APRFKQILLDKREIDFNVLVEYFN-PLS 31
DB 121 APRFKQVLSDLQRLDSHTLIMGDFNTPLS 151

RESULT 5
O00366 PRELIMINARY; PRT; 1275 AA.
AC O00366;
DT 01-JUL-1997 (TRENBLrel. 04, Created)
DT 01-JUL-1997 (TRENBLrel. 04, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Putative p150.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97285120; PubMed=9140393;
RA Sasaman D.M., Dombroski B.A., Moran J.V., Kimberland M.L., Naas T.P.,
RA DeBerardinis R.J., Gabriel A., Swergold G.D., Kazazian H.H. Jr.;
RT "Many human Li elements are capable of retrotransposition.";
RL Nat. Genet. 16:37-43(1997).
DR EMBL; U93567; AAC51267.1; -.
DR PIR; B28096; B28096.
DR PIR; S23650; S23650.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006278; F:RNA-dependent DNA replication; IEA.
DR InterPro; IPR005135; Exo_endo_phos.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF03372; Exo_endo_phos; 1.
DR Pfam; PF00078; RVT; 1.
KW RNA-directed DNA polymerase; Transferase.
SQ SEQUENCE 1275 AA; 149033 MW; 07E88F8F4DB831A2 CRC64;

Query Match 50.3%; Score 78.5; DB 2; Length 1275;
Best Local Similarity 54.8%; Pred. No. 0.032;

```

	Matches	17;	Conservative	7;	Mismatches	6;	Indels	1;	Gaps	1;
Qy	2	APRFKQILLDLKREIDFNRLVEYFN-PLS 31								
		:: :::	:	:	:	:	:	:	:	:
Dd	121	APRFKQVLDLQRLDSHTLMGDFNTPLS 151								
RESULT 6										
ID	O00368	PRELIMINARY;	PRT;	1275	AA.					
AC	O00368;									
DT	01-JUL-1997	(TREMBLrel. 04, Created)								
DT	01-JUL-1997	(TREMBLrel. 04, Last sequence update)								
DT	01-OCT-2003	(TREMBLrel. 25, Last annotation update)								
DE	Putative p150.									
OS	Homo sapiens (Human).									
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.									
CX	NCBI_TaxID=9606;									
[1]										
RN	SEQUENCE FROM N.A.									
RP	MEDLINE=97285120; PubMed=9140393;									
RX	Sassaman D.M., Domroski B.A., Moran J.V., Kimberland M.L., Naas T.P.,									
RA	DeBerardinis R.J., Gabriel A., Swergold G.D., Kazanian H.H. Jr.;									
RA	"Many human L1 elements are capable of retrotransposition.";									
RL	Nat. Genet. 16:137-43(1997).									
DR	EMBL; U93568; AAC51269.1; ..									
DR	PIR; B28096; B28096.									
DR	PIR; S23650; S23650.									
DR	GO; GO:0003723; F:RNA binding; IEA.									
DR	GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.									
DR	GO; GO:0006278; P:RNA-dependent DNA replication; IEA.									
DR	InterPro; IPR005135; Exo_endo_phos.									
DR	InterPro; IPR000477; RVTse.									
DR	Pfam; PF03372; Exo_endo_phos; 1.									
DR	Pfam; PF00078; RVT; 1.									
SK	RNA-directed DNA polymerase; Transferase.									
QW	SEQUENCE 1275 AA; 149077 MW; 3BEC3E2DC2E0EB61 CRC64;									
	Query Match	50.3%;	Score	78.5;	DB 2;	Length	1275;			
	Best Local Similarity	54.8%;	Pred. No.	0.032;						
Matches	17;	Conservative	7;	Mismatches	6;	Indels	1;	Gaps	1;	
Qy	2	APRFKQILLDLKREIDFNRLVEYFN-PLS 31								
		:: :::	:	:	:	:	:	:	:	:
Dd	121	APRFKQVLDLQRLDSHTLMGDFNTPLS 151								
RESULT 7										
ID	O00370	PRELIMINARY;	PRT;	1275	AA.					
AC	O00370;									
DT	01-JUL-1997	(TREMBLrel. 04, Created)								
DT	01-JUL-1997	(TREMBLrel. 04, Last sequence update)								
DT	01-OCT-2003	(TREMBLrel. 25, Last annotation update)								
DE	Putative p150.									
OS	Homo sapiens (Human).									
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.									
CX	NCBI_TaxID=9606;									
[1]										
RN	SEQUENCE FROM N.A.									
RP	MEDLINE=97285120; PubMed=9140393;									
RA	Sassaman D.M., Domroski B.A., Moran J.V., Kimberland M.L., Naas T.P.,									
RA	DeBerardinis R.J., Gabriel A., Swergold G.D., Kazanian H.H. Jr.;									
RA	"Many human L1 elements are capable of retrotransposition.";									
RL	Nat. Genet. 16:137-43(1997).									
DR	EMBL; U93569; AAC51271.1; ..									
DR	PIR; B28096; B28096.									

RA DeBerardinis R.J., Gabriel A., Swergold G.D., Kazazian H.H. Jr.;
 RT "Many human l1 elements are capable of retrotransposition.";
 RL Nat. Genet. 16:37-43(1997).
 DR EMBL: U93574; AAC51279.1; ..
 DR PIR: B28096; B28096.
 DR PIR: J00033; J00033.
 DR PIR: S23650; S23650.
 DR GO: GO:0003723; F:RNA binding; IEA.
 DR GO: GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
 DR GO: GO:0006278; P:RNA-dependent DNA replication; IEA.
 DR InterPro: IPR005135; Exo_endo_phos.
 DR InterPro: IPR000477; RVTse.
 DR Pfam: PF03372; Exo_endo_phos; 1.
 DR Pfam: PF00078; RVT; 1.
 DR RNA-directed DNA polymerase; Transferase.
 KW RNA-directed DNA polymerase; Transferase.
 SQ SEQUENCE 1275 AA; 148879 MW; F006F74BBB72B87 CRC64;

Query Match 50.3%; Score 78.5; DB 2; Length 1275;
 Best Local Similarity 54.8%; Pred. No. 0.032; Indels 1; Gaps 1;
 Matches 17; Conservative 7; Mismatches 6; Indels 1; Gaps 1;
 QY 2 APRFIKQILLDLKREIDFNRLVEYFN-PLS 31
 DB 121 APRFIKQVLSLDRLDLSHTLMGDFNTPLS 151

RESULT 10
 QSTE30 PRELIMINARY; PRT; 1275 AA.
 ID QSTE30;
 AC QSTE30;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 ON NCBI_TaxID=9606;
 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21668188; PubMed=11810275;
 RA Benjes S.W., Morris C.W.;
 RT "A full-length and potentially active LINE element is integrated
 RT polymorphically within the IGL locus in a genomically unstable region
 RT of chromosome 22.";
 RL Hum. Genet. 109:628-637(2001).
 DR EMBL: AF421375; AAL50637.1; ..
 DR PIR: B28096; B28096.
 DR PIR: S23650; S23650.
 DR GO: GO:0003723; F:RNA binding; IEA.
 DR GO: GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
 DR GO: GO:0006278; P:RNA-dependent DNA replication; IEA.
 DR InterPro: IPR005135; Exo_endo_phos.
 DR InterPro: IPR000477; RVTse.
 DR Pfam: PF03372; Exo_endo_phos; 1.
 DR Pfam: PF00078; RVT; 1.
 DR RNA-directed DNA polymerase; Transferase.
 KW Hypothetical protein; RNA-directed DNA polymerase; Transferase.
 SQ SEQUENCE 1275 AA; 149009 MW; B327D9D50A581764 CRC64;

Query Match 50.3%; Score 78.5; DB 2; Length 1275;
 Best Local Similarity 54.8%; Pred. No. 0.032; Indels 1; Gaps 1;
 Matches 17; Conservative 7; Mismatches 6; Indels 1; Gaps 1;
 QY 2 APRFIKQILLDLKREIDFNRLVEYFN-PLS 31
 DB 121 APRFIKQVLSLDRLDLSHTLMGDFNTPLS 151

RESULT 11
 Q6ZN98 PRELIMINARY; PRT; 244 AA.
 ID Q6ZN98
 AC Q6ZN98;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)

DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein FLJ16301.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 ON NCBI_TaxID=9606;
 [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RA Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
 RA Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,
 RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
 RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
 RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
 RA Murakawa K., Kanetori K., Takahashi-Fujii A., Oshima A., Sugiyama A.,
 RA Kawakami B., Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K.,
 RA Isonai T.;
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AK131313; BAD18477.1; ..
 DR InterPro: IPR005135; Exo_endo_phos.
 DR Pfam: PF03372; Exo_endo_phos; 1.
 SQ SEQUENCE 244 AA; 28032 MW; 0B7AD3D0357935B1 CRC64;

Query Match 50.0%; Score 78; DB 2; Length 244;
 Best Local Similarity 53.6%; Pred. No. 0.0064;
 Matches 15; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 1 RAPRFIKQILLDLKREIDFNRLVEYFN 28
 DB 119 RAPRFIKQVLRDLRLDLSHTLMGDFN 146

RESULT 12
 BAD18477 PRELIMINARY; PRT; 244 AA.
 ID BAD18477;
 AC BAD18477;
 DT 12-MAY-2004 (TrEMBLrel. 27, Created)
 DT 12-MAY-2004 (TrEMBLrel. 27, Last sequence update)
 DT 12-MAY-2004 (TrEMBLrel. 27, Last annotation update)
 DE CDNA FLJ16301 fis, clone PLACE7000333.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 ON NCBI_TaxID=9606;
 [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RA Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
 RA Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,
 RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
 RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
 RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
 RA Murakawa K., Kanetori K., Takahashi-Fujii A., Oshima A., Sugiyama A.,
 RA Kawakami B., Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K.,
 RA Isonai T.;
 RL "NEBO human cDNA sequencing project";
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AK131313; BAD18477.1; ..
 SQ SEQUENCE 244 AA; 28032 MW; 0B7AD3D0357935B1 CRC64;

Query Match 50.0%; Score 78; DB 2; Length 244;
 Best Local Similarity 53.6%; Pred. No. 0.0064;
 Matches 15; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 1 RAPRFIKQILLDLKREIDFNRLVEYFN 28
 DB 119 RAPRFIKQVLRDLRLDLSHTLMGDFN 146

RESULT 13
 L1N1_HUMAN STANDARD; PRT; 1259 AA.
 ID L1N1_HUMAN


```

RESULT 15
BADIID BADI8452 PRELIMINARY; PRT; 314 AA.
AC BADI8452;
DDT 12-MAY-2004 (TrEMBLrel. 27, Created)
DDT 12-MAY-2004 (TrEMBLrel. 27, Last sequence update)
DDT 12-MAY-2004 (TrEMBLrel. 27, Last annotation update)
DE CDNA FJ16220 f1s, clone CTRONG3002552. (Fragment).
DOS Homo sapiens (Human).
OCC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OCC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI TaxID=9606;
[1]
NRN SEQUENCE FROM N.A.
RRP TISSUE=Tongue;
RCR Kanehori K., Ishibashi T., Chiba Y., Fujimori K., Hiraoaka S.,
RA Tanai H., Watanabe S., Ishida S., Ono Y., Hotuta T., Watanabe M.,
RA Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S.,
RA Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA Wagatsuma M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahari K., Masuko Y., Nagai K., Isogai T.;
RA "NEDO human cDNA sequencing project.";
RT Submitted (MAR-2004) to the EMBL/GenBank/DBSJ databases.
REL EMBL; AK131275; BADI8452.1; -.
DR NON TER 314 314
FTS SEQUENCE 314 AA; 36311 MW; 2C37DD7DD1001A56 CRC64;

Query Match 48.4%; Score 75.5; DB 2; Length 314;
Best Local Similarity 51.6%; Pred. No. 0.019;
Matches 16; Conservative 8; Mismatches 6; Indels 1; Gaps 1;

QY 2 APRRIKQLLDLRKREIDFNVLVEYN-PLS 31
||| ||| : | : : : : |||
Ddb 121 APRRIKVLSDVQRDLDSHTLINGDFNTPLS 151

Search completed: November 10, 2004, 13:38:53
Job time : 39.8532 secs

```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:44:14 ; Search time 9.51698 Seconds
(without alignments)
181.178 Million cell updates/sec

Title: US-10-092-750-44

Perfect score: 125

Sequence: 1 IVALIAGRLRLMLGDFNGELEASAKN 26

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A.COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B.COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PTUTS.COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52	41.6	3519	3	US-09-428-517-4
2	49	39.2	3816	3	US-09-428-517-3
3	49	39.2	4150	3	US-09-428-517-2
4	44.5	35.6	40	3	US-08-630-915A-82
5	44.5	35.6	40	4	US-09-879-957-82
6	44.5	35.6	351	4	US-09-134-000C-5390
7	44	35.2	268	4	US-09-134-000C-5975
8	44	35.2	315	4	US-09-252-991A-27945
9	44	35.2	415	4	US-09-252-991A-23751
10	43.5	34.8	48	4	US-09-023-905A-16
11	43.5	34.8	54	1	US-08-167-035-28
12	43.5	34.8	54	1	US-08-208-887A-28
13	43.5	34.8	54	2	US-08-539-005-28
14	43.5	34.8	54	3	US-09-346-510B-19
15	43.5	34.8	54	3	US-09-280-598-30
16	43.5	34.8	348	4	US-09-107-532A-5421
17	43.5	34.8	800	6	5183745-3
18	43.5	34.8	1130	4	US-09-538-092-834
19	43	34.4	1025	2	US-08-530-792D-23
20	43	34.4	1026	2	US-08-530-792D-22
21	42.5	34.0	271	4	US-09-252-991A-26051
22	42	33.6	161	4	US-09-544-664B-46
23	42	33.6	27	4	US-09-544-664B-19
24	42	33.6	161	4	US-10-101-464A-729
25	42	33.6	349	4	US-09-252-991A-25984
26	42	33.6	502	4	US-09-252-991A-27619
27	42	33.6	530	3	US-09-199-637A-130

RESULT 1

US-09-428-517-4
; Sequence 4, Application US/09428517
; Patent No. 6251636
; GENERAL INFORMATION:
; APPLICANT: Betlach, Mary C.
; APPLICANT: Shah, Sanjay Krishnakant
; APPLICANT: McDaniel, Robert
; APPLICANT: Tang, Li
; TITLE OF INVENTION: RECOMBINANT OLEANDOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 30062-20039.00
; CURRENT APPLICATION NUMBER: US/09/428,517
; CURRENT FILING DATE: 1999-10-28
; EARLIER APPLICATION NUMBER: 60/120,254
; EARLIER FILING DATE: 1999-02-16
; EARLIER APPLICATION NUMBER: 60/106,100
; EARLIER FILING DATE: 1998-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 3519
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Recombinant
; OTHER INFORMATION: Oleandolide PKS
US-09-428-517-4

Query Match 41.6%; Score 52; DB 3; Length 3519;
Best Local Similarity 40.0%; Pred. No. 19;
Matches 10; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 2 VAIAGRLRLMLGDFNGELEASAKN 26

DB 701 VSLPAGRVMTMLEFDGRLSVAVN 725

RESULT 2

US-09-428-517-3
; Sequence 3, Application US/09428517
; Patent No. 6251636
; GENERAL INFORMATION:
; APPLICANT: Betlach, Mary C.
; APPLICANT: Shah, Sanjay Krishnakant
; APPLICANT: McDaniel, Robert
; APPLICANT: Tang, Li
; TITLE OF INVENTION: RECOMBINANT OLEANDOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 30062-20029.00
; CURRENT APPLICATION NUMBER: US/09/428,517
; CURRENT FILING DATE: 1999-10-28

Sequence 29592, A
Sequence 3906, Ap
Sequence 2, Appli
Sequence 6, Appli
Sequence 47579, A
Sequence 33163, A
Sequence 5770, Ap
Sequence 46, Appli
Sequence 2, Appli
Sequence 21011, A
Sequence 26045, A
Sequence 24230, A
Sequence 2465, Ap
Sequence 24047, A
Sequence 3, Appli
Sequence 6, Appli
Sequence 3, Appli
Sequence 2, Appli

ALIGNMENTS

```

; EARLIER APPLICATION NUMBER: 60/120,254
; EARLIER FILING DATE: 1999-02-16
; EARLIER APPLICATION NUMBER: 60/106,100
; EARLIER FILING DATE: 1998-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 3
; LENGTH: 3816
;

```

Query Match 39.2%; Score 49; DB 3; Length 3816;
Best Local Similarity 40.0%; Pred. No. 67;
Matches 10; Conservative 6; Mismatches 9; Indels

```

RESULT 3
US-09-428-517-2
; Sequence 2, Application US/09428517
; Patent No. 6251636
; GENERAL INFORMATION:
; APPLICANT: Betlach, Mary C.
; APPLICANT: Shah, Sanjay Krishnakant
; APPLICANT: McDaniel, Robert
; APPLICANT: Tang, Li
; TITLE OF INVENTION: RECOMBINANT OLEANDOLIDE POLYPEPTIDE SYNTHASE
; FILE REFERENCE: 30062-20029.00
; CURRENT APPLICATION NUMBER: US/09/428,517
; CURRENT FILING DATE: 1999-10-28
; EARLIER APPLICATION NUMBER: 60/120,254
; EARLIER FILING DATE: 1999-02-16
; EARLIER APPLICATION NUMBER: 60/106,100
; EARLIER FILING DATE: 1998-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4150
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Recombinant
; OTHER INFORMATION: Oleandolide PKS
US-09-428-517-2

```

Query Match	39.2%	Score 49;	DB 3;	Length 4150;
Best Local Similarity	36.0%;	Pred. No. 74;		
Matches 9;	Conservative 9;	Mismatches 7;	Indels 0;	Gaps 0;

RESULT 4
US-08-630-915A-82
Sequence 82, Application US/08630915A
Patent No. 630820
GENERAL INFORMATION:
APPLICANT: SPARKS, Andrew B.
APPLICANT: HOFFMAN, No. 6309820h
APPLICANT: KAT, Brian K.
APPLICANT: FOMKES, Dana M.
APPLICANT: MCCONNELL, Stephen J.
TITLE OF INVENTION: POLYPEPTIDES HAVING
TITLE OF INVENTION: DOMAIN OF INTEREST

1 TITLE OF INVENTION: USING SAME
 2
 3 NUMBER OF SEQUENCES: 227
 4
 5 CORRESPONDENCE ADDRESS:
 6
 7 ADDRESSSEE: Pennie & Edmonds LLP
 8
 9 STREET: 1155 Avenue of the Americas
 10
 11 CITY: New York
 12
 13 STATE: New York
 14
 15 COUNTRY: USA
 16
 17 ZIP: 10036-2711
 18
 19 COMPUTER READABLE FORM:
 20
 21 MEDIUM TYPE: Floppy disk
 22
 23 COMPUTER: IBM PC compatible
 24
 25 OPERATING SYSTEM: PC-DOS/MS-DOS
 26
 27 SOFTWARE: PatentIn Release #1.0, Version #1.30
 28
 29 CURRENT APPLICATION DATA:
 30
 31 APPLICATION NUMBER: US/08/630.915A
 32
 33 FILING DATE: 03-APR-1996
 34
 35 CLASSIFICATION: 536

Query Match	35.8%	Score 44.5;	DB 3;	Length 40;
Best Local Similarity	57.3%	Pred. No. 1.6;		
Matches 11;	Conservative	2;	Mismatches	5;
			Indels	1;
			Caps	1;

```

RESULT 5
US-09-879-957-82
; Sequence 82, Application US/09879957
; Patent No. 6709821
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; HOFFMAN, NO. 6709821h
; KAY, Brian K.
; FOWLKES, Dana M.
; MCCONNELL, Stephen J.
; TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
; DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
; USING SAME
;
;

```

```

1 OPERATING SYSTEM: PC-DOS/MS-DOS
2 SOFTWARE: PatentIn Release #1.0, Version #1.30
3
4 CURRENT APPLICATION DATA: US/09/879,957
5 APPLICATION NUMBER: US/09/879,957
6 FILING DATE: 13-Jun-2001
7

```

```
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 08/630,915
;   FILING DATE: 03-APR-1996
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Misrock, S. Leslie
;     REGISTRATION NUMBER: 18,872
;     REFERENCE/DOCKET NUMBER: 1101-174
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (212) 790-9090
;   TELEFAX: (212) 869-8864/9741
;   TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 82:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 40 amino acids
;     TYPE: amino acid
;     STRANDEDNESS: <Unknown>
;     TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 82:
US-09-879-957-82

Query Match      35.6%; Score 44.5; DB 4; Length 40;
Best Local Similarity 57.9%; Pred. No. 1.6;
Matches 11; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY      8 RLRLMGDFNGELESAXN 26
      :|||:|||||:
Db      21 KRLVGYNHNGE-EAQTKN 38

RESULT 6
US-09-134-000C-5390
; Sequence 5390, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5390
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-5390

Query Match      35.6%; Score 44.5; DB 4; Length 351;
Best Local Similarity 40.7%; Pred. No. 22;
Matches 11; Conservative 6; Mismatches 7; Indels 3; Gaps 1;

QY      1 IVVAIAGRLRLM---GDFNGELESASA 24
      :|||:|||||:
Db      258 IVAVGTTTSIRLTLEIGTKFNGEIQADS 284

RESULT 7
US-09-134-000C-5975
; Sequence 5975, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778

; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23751
; LENGTH: 415
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23751

Query Match      35.2%; Score 44; DB 4; Length 315;
Best Local Similarity 45.8%; Pred. No. 23;
Matches 11; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY      3 AIIAGRLRLMGDFNGELESASAXN 26
      :|||:|||||:
Db      285 AIVAGRLAYLAGMPKLYASASS 308

RESULT 9
US-09-252-991A-23751
; Sequence 23751, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23751
; LENGTH: 415
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23751

Query Match      35.2%; Score 44; DB 4; Length 268;
Best Local Similarity 50.0%; Pred. No. 19;
Matches 11; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY      1 IVVAIAGRLRLMGDFNGELEA 22
      :|||:|||||:
Db      91 LVYVIDGLRLVSDGQETHELEA 112

RESULT 8
US-09-252-991A-27945
; Sequence 27945, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27945
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27945

Query Match      35.2%; Score 44; DB 4; Length 315;
Best Local Similarity 45.8%; Pred. No. 23;
Matches 11; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY      3 AIIAGRLRLMGDFNGELESASAXN 26
      :|||:|||||:
Db      285 AIVAGRLAYLAGMPKLYASASS 308
```

Matches 10; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 2 VAILAGRLMLGDPNGLEASA 24
DB 142 VALLRGRLPLHDHFGRRDHRA 164

RESULT 10
US-09-023-905A-16
; Sequence 16, Application US/09023905A
; Patent No. 6475778
; GENERAL INFORMATION:
; APPLICANT: Roberts, Thomas M.
; APPLICANT: King, Frederick J.
; APPLICANT: Harris, David F.
; APPLICANT: Hu, Erding
; APPLICANT: Spiegelman, Bruce
; APPLICANT: Chan, Joanne
; TITLE OF INVENTION: Differentiation Enhancing Factors and Uses
; FILE REFERENCE: DFN-021
; CURRENT APPLICATION NUMBER: US/09/023,905A
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/038,191
; PRIOR FILING DATE: 1997-02-14
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-023-905A-16

Query Match 34.8%; Score 43.5; DB 4; Length 48;
Best Local Similarity 55.0%; Pred. No. 2.9;
Matches 11; Conservative 2; Mismatches 6; Indels 1; Gaps 1;

QY 8 RLRLMGDPNGE-LEASAKN 26
DB 20 KLRLVLYNHNGEWCEAQTKN 39

RESULT 11
US-08-167-035-28
; Sequence 28, Application US/08167035
; Patent No. 5618691
; GENERAL INFORMATION:
; APPLICANT: Schlessinger, Joseph
; APPLICANT: Skolnick, Edward Y.
; APPLICANT: Margolis, Benjamin L.
; TITLE OF INVENTION: NOVEL EXPRESSION CLONING METHOD FOR
; TITLE OF INVENTION: IDENTIFYING TARGET PROTEINS FOR EUKARYOTIC TYROSINE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: 10036-2711
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/167,035
; FILING DATE: 16-DEC-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742

; REFERENCE/DOCKET NUMBER: 7683-062
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 54 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-167-035-28

Query Match 34.8%; Score 43.5; DB 1; Length 54;
Best Local Similarity 55.0%; Pred. No. 3.4;
Matches 11; Conservative 2; Mismatches 6; Indels 1; Gaps 1;

QY 8 RLRLMGDPNGE-LEASAKN 26
DB 20 KLRLVLYNHNGEWCEAQTKN 39

RESULT 12
US-08-208-887A-28
; Sequence 28, Application US/08208887A
; Patent No. 5677421
; GENERAL INFORMATION:
; APPLICANT: Schlessinger, Joseph
; APPLICANT: Skolnick, Edward Y.
; APPLICANT: Margolis, Benjamin L.
; TITLE OF INVENTION: NOVEL EXPRESSION CLONING METHOD FOR
; TITLE OF INVENTION: IDENTIFYING TARGET PROTEINS FOR EUKARYOTIC TYROSINE
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: 10036-2711
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/208,887A
; FILING DATE: 11-MAR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7683-063
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 54 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-208-887A-28

Query Match 34.8%; Score 43.5; DB 1; Length 54;
Best Local Similarity 55.0%; Pred. No. 3.4;
Matches 11; Conservative 2; Mismatches 6; Indels 1; Gaps 1;

QY 8 RLRLMGDPNGE-LEASAKN 26
DB 20 KLRLVLYNHNGEWCEAQTKN 39

```
RESULT 13
US-08-539-005-28
; Sequence 28, Application US/08539005
; Patent No. 5858686
; GENERAL INFORMATION:
; APPLICANT: Schllessinger, Joseph
; APPLICANT: Skolnick, Edward Y.
; APPLICANT: Margolis, Benjamin L.
; TITLE OF INVENTION: NOVEL EXPRESSION CLONING METHOD FOR
; TITLE OF INVENTION: IDENTIFYING TARGET PROTEINS FOR EUKARYOTIC TYROSINE
; TITLE OF INVENTION: KINASES AND NOVEL TARGET PROTEINS
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: 10036-2711
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/539,005
; FILING DATE: 4-OCT-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/167,035
; FILING DATE: 16-DEC-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7683-062
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 54 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-539-005-28

Query Match 34.8%; Score 43.5; DB 2; Length 54;
Best Local Similarity 55.0%; Pred. No. 3.4;
Matches 11; Conservative 2; Mismatches 6; Indels 1; Gaps 1;

QY 8 RLRLMGDQFNCE-LEASAKN 26
Db 20 KLRLVLYNHNGWCEAQTKN 39

RESULT 14
US-09-346-510B-19
; Sequence 19, Application US/09346510B
; Patent No. 6281014
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Wang, Yinxiang
; TITLE OF INVENTION: SH3-Containing Protein, DNA and Uses Thereof
; FILE REFERENCE: D6221C1P
; CURRENT APPLICATION NUMBER: US/09/346,510B
; CURRENT FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: 08/871,732
; PRIOR FILING DATE: 1997-06-09
; NUMBER OF SEQ ID NOS: 32
```

```
; SEQ ID NO 19
; LENGTH: 54
; TYPE: PRT
; ORGANISM: unknown
; FEATURE:
; NAME/KEY: Domain
; OTHER INFORMATION: amino acid sequence of C-Abl SH3 domain
US-09-346-510B-19

Query Match 34.8%; Score 43.5; DB 3; Length 54;
Best Local Similarity 55.0%; Pred. No. 3.4;
Matches 11; Conservative 2; Mismatches 6; Indels 1; Gaps 1;

QY 8 RLRLMGDQFNCE-LEASAKN 26
Db 20 KLRLVLYNHNGWCEAQTKN 39

RESULT 15
US-09-280-598-30
; Sequence 30, Application US/09280598
; Patent No. 6391584
; GENERAL INFORMATION:
; APPLICANT: Schllessinger, Joseph
; APPLICANT: Skolnick, Edward Y.
; APPLICANT: Margolis, Benjamin L.
; APPLICANT: App, Harold
; TITLE OF INVENTION: A NOVEL EXPRESSION-CLONING METHOD FOR
; TITLE OF INVENTION: IDENTIFYING TARGET PROTEINS FOR EUKARYOTIC TYROSINE
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/280,598
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/252,820
; FILING DATE: 02-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7683-067
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 54 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-09-280-598-30

Query Match 34.8%; Score 43.5; DB 3; Length 54;
Best Local Similarity 55.0%; Pred. No. 3.4;
Matches 11; Conservative 2; Mismatches 6; Indels 1; Gaps 1;

QY 8 RLRLMGDQFNCE-LEASAKN 26
Db 20 KLRLVLYNHNGWCEAQTKN 39
```

Fri Nov 12 14:55:27 2004

us-10-092-750-44.ra1

Page 6

Search completed: November 10, 2004, 14:55:33
Job time : 9.56698 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2004, 01:28:30 ; Search time 30.0717 Seconds
(without alignments)
305.399 Million cell updates/sec

Title: US-10-092-750-44
Perfect score: 125
Sequence: 1 IVALLAGRLMLGDFNGELEASAKN 26

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	125	100.0	26	14 US-10-092-750-44	Sequence 44, Appl
2	88	70.4	71	17 US-10-425-115-238718	Sequence 238718,
3	52	41.6	363	9 US-09-925-300-1530	Sequence 1530, Ap
4	52	41.6	1072	15 US-10-307-817-90	Sequence 90, Appl
5	52	41.6	1072	15 US-10-287-226-230	Sequence 230, App
6	52	41.6	1091	15 US-10-307-817-469	Sequence 469, App
7	52	41.6	1091	15 US-10-287-226-669	Sequence 669, App
8	52	41.6	1095	15 US-10-307-817-86	Sequence 86, Appl
9	52	41.6	1095	15 US-10-287-226-232	Sequence 232, App
10	52	41.6	1095	15 US-10-287-226-238	Sequence 238, App
11	52	41.6	1095	15 US-10-287-226-246	Sequence 246, App
12	52	41.6	1100	15 US-10-287-226-668	Sequence 668, App
13	52	41.6	1101	15 US-10-307-817-84	Sequence 84, Appl

14 52 41.6 1101 15 US-10-307-817-92 Sequence 92, Appl
15 52 41.6 1101 15 US-10-287-226-226 Sequence 226, App
16 52 41.6 1101 15 US-10-287-226-228 Sequence 228, App
17 52 41.6 1101 15 US-10-287-226-254 Sequence 254, App
18 52 41.6 1101 15 US-10-287-226-651 Sequence 651, App
19 52 41.6 1102 15 US-10-287-226-242 Sequence 242, App
20 52 41.6 1103 15 US-10-287-226-252 Sequence 252, App
21 52 41.6 1106 15 US-10-287-226-236 Sequence 236, App
22 52 41.6 1106 15 US-10-287-226-244 Sequence 244, App
23 52 41.6 1116 15 US-10-287-226-250 Sequence 250, App
24 52 41.6 3519 10 US-09-808-880-4 Sequence 13359, A
25 49 39.2 220 14 US-10-369-493-13359 Sequence 279915, A
26 49 39.2 307 17 US-10-435-115-279915 Sequence 49176, A
27 49 39.2 308 15 US-10-435-114-49176 Sequence 2, Appl
28 49 39.2 3816 10 US-09-808-880-3 Sequence 2, Appl
29 49 39.2 4150 10 US-09-808-880-2 Sequence 238650,
30 47 37.6 150 15 US-10-424-599-238650 Sequence 221584,
31 47 37.6 152 17 US-10-435-115-221584 Sequence 298169,
32 47 37.6 843 17 US-10-425-115-298169 Sequence 281530,
33 47 37.6 1544 17 US-10-435-115-281530 Sequence 46728, A
34 46 36.8 722 16 US-10-437-963-176171 Sequence 176171, A
35 46 36.8 722 16 US-10-369-493-10870 Sequence 10870, A
36 45 36.0 476 14 US-10-156-761-15070 Sequence 15070, A
37 45 36.0 523 14 US-10-156-761-15070 Sequence 66108, A
38 45 36.0 2703 15 US-10-282-122A-66108 Sequence 65564, A
39 45 36.0 2799 15 US-10-282-122A-65564 Sequence 82, Appl
40 44.5 35.6 40 9 US-09-879-957-82 Sequence 114869,
41 44.5 35.6 40 16 US-10-807-856-82 Sequence 10781, A
42 44.5 35.6 216 16 US-10-437-963-114868 Sequence 56946, A
43 44.5 35.6 346 9 US-09-815-242-10781 Sequence 11278, A
44 44.5 35.6 346 15 US-10-282-122A-56946
45 44.5 35.6 573 14 US-10-369-493-11278

ALIGNMENTS

RESULT 1
US-10-092-750-44
; Sequence 44, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; APPLICANT: Wright, Martin C.
; TITLE OF INVENTION: Polypeptides Interactive with BCL-XL
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092,750
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/274,526
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-44

Query Match 100.0%; Score 125; DB 14; Length 26;
Best Local Similarity 100.0%; Pred. No. 3,le-13;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IVALLAGRLMLGDFNGELEASAKN 26
DB 1 IVALLAGRLMLGDFNGELEASAKN 26

RESULT 2
US-10-425-115-238718
; Sequence 238718, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:

```

; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 238718
; LENGTH: 71
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_14929C.1.pap
US-10-425-115-238718

Query Match      70.4%; Score 88; DB 17; Length 71;
Best Local Similarity 89.5%; Pred. No. 1.2e-06;
Matches 17; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAIAGRLRLMLGDPNGEL 20
DB 53 VAIVAGRLRLMLGDKFNGEL 71

RESULT 3
US-09-925-300-1530
; Sequence 1530, Application US/09925300
; Publication No. US2002015161A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1530
; LENGTH: 363
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (178)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (179)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1530

Query Match      41.6%; Score 52; DB 9; Length 363;
Best Local Similarity 40.0%; Pred. No. 6.7;
Matches 10; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 IVAIAGRLRLMLGDPNGELEASAK 25
DB 182 LVSSLTSGLLTIGDRFGALDAAK 206

RESULT 4
US-10-307-817-90
; Sequence 90, Application US/10307817
; Publication No. US2004005838A1
; GENERAL INFORMATION:
; APPLICANT: Agee et al.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME

```

```

; FILE REFERENCE: 21402-502C
; CURRENT APPLICATION NUMBER: US/10/307,817
; CURRENT FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 682
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 90
; LENGTH: 1072
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-307-817-90

Query Match      41.6%; Score 52; DB 15; Length 1072;
Best Local Similarity 40.0%; Pred. No. 24;
Matches 10; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 IVAIAGRLRLMLGDPNGELEASAK 25
DB 891 LVSSLTSGLLTIGDRFGALDAAK 915

RESULT 5
US-10-287-226-230
; Sequence 230, Application US/10287226
; Publication No. US20040086875A1
; GENERAL INFORMATION:
; APPLICANT: Agee, Michele L.,
; APPLICANT: Alsbrook, John P.,
; APPLICANT: Berghs, Constance,
; APPLICANT: Boldog, Ference,
; APPLICANT: Burgess, Catherine E.,
; APPLICANT: Chant, John S.,
; APPLICANT: Chaudhuri, Amitabha,
; APPLICANT: DiPippo, Vincent A.,
; APPLICANT: Edinger, Shlomit R.,
; APPLICANT: Eisen, Andrew,
; APPLICANT: Ellerman, Karen,
; APPLICANT: Gangolli, Esha A.,
; APPLICANT: Gorman, Linda,
; APPLICANT: Gerlach, Valerie,
; APPLICANT: Ji, Weizhen,
; APPLICANT: Kekuda, Ramesh,
; APPLICANT: Khramtsov, Nikolai,
; APPLICANT: Li, Li,
; APPLICANT: Malyankar, Uriel M.,
; APPLICANT: MacDougall, John R.,
; APPLICANT: Mezes, Peter S.
; APPLICANT: Miller, Charles E.,
; APPLICANT: Millet, Isabelle,
; APPLICANT: Ooi, Chean Eng,
; APPLICANT: Ort, Tatiana,
; APPLICANT: Padigaru, Muralidhara,
; APPLICANT: Patturajan, Meera,
; APPLICANT: Rastelli, Luca,
; APPLICANT: Rieger, Daniel K.,
; APPLICANT: Rothenberg, Mark E.,
; APPLICANT: Shenoy, Suresh G.,
; APPLICANT: Spaderina, Steven K.,
; APPLICANT: Spytek, Kimberley A.,
; APPLICANT: Taupier, Jr., Raymond J.,
; APPLICANT: Vernet, Corine A.M.,
; APPLICANT: Zerhusen, Bryan D.,
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-480C
; CURRENT APPLICATION NUMBER: US/10/287,226
; CURRENT FILING DATE: 2002-11-04
; PRIOR APPLICATION NUMBER: 60/334,421
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/354,392
; PRIOR FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: 60/360,148
; PRIOR FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: 60/364,000

```

```
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 60/404,821
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: 60/334,526
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/354,409
; PRIOR FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: 60/364,227
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 60/334,027
; PRIOR FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: 60/331,641
; PRIOR FILING DATE: 2001-11-20
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 673
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 230
; LENGTH: 1072
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-287-226-230

Query Match 41.6%; Score 52; DB 15; Length 1072;
Best Local Similarity 40.0%; Pred. No. 24;
Matches 10; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 IVAIAGRLRLMGDPNGELEASEAK 25
Db 891 LVSLTSGLLTIGDRFGGALDAAK 915

RESULT 6
US-10-307-817-469
; Sequence 469, Application US/10307817
; Publication No. US20040058338A1
; GENERAL INFORMATION:
; APPLICANT: Agee et al.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-502C
; CURRENT APPLICATION NUMBER: US/10/307,817
; CURRENT FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 682
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 469
; LENGTH: 1091
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-307-817-469

Query Match 41.6%; Score 52; DB 15; Length 1091;
Best Local Similarity 40.0%; Pred. No. 24;
Matches 10; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 IVAIAGRLRLMGDPNGELEASEAK 25
Db 910 LVSLTSGLLTIGDRFGGALDAAK 934

RESULT 7
US-10-287-226-669
; Sequence 669, Application US/10287226
; Publication No. US20040086875A1
; GENERAL INFORMATION:
; APPLICANT: Agee, Michele L.,
; APPLICANT: Alsobrook, John P.,
; APPLICANT: Berghs, Constance,
; APPLICANT: Boldog, Ference,
; APPLICANT: Burgess, Catherine E.,
; APPLICANT: Chant, John S.,
; APPLICANT: Chaudhuri, Anitabha,
; APPLICANT: DiPippo, Vincent A.,
; APPLICANT: Edinger, Shommit R.,
; APPLICANT: Eisen, Andrew,
```

```
; APPLICANT: Ellerman, Karen,
; APPLICANT: Gangolli, Esha A.,
; APPLICANT: Gorman, Linda,
; APPLICANT: Gerlach, Valerie,
; APPLICANT: Ji, Weizhen,
; APPLICANT: Kekuda, Ramesh,
; APPLICANT: Khrantsov, Nikolai,
; APPLICANT: Li, Li,
; APPLICANT: Malyankar, Uriel M.,
; APPLICANT: Macbougall, John R.,
; APPLICANT: Mezes, Peter S.
; APPLICANT: Miller, Charles E.,
; APPLICANT: Millet, Isabelle,
; APPLICANT: Ooi, Chean Eng,
; APPLICANT: Ort, Tatiana,
; APPLICANT: Padigar, Muralidhara,
; APPLICANT: Patturajan, Meera,
; APPLICANT: Rastelli, Luca,
; APPLICANT: Rieger, Daniel K.,
; APPLICANT: Rothenberg, Mark E.,
; APPLICANT: Shenoy, Suresh G.,
; APPLICANT: Spaderna, Steven K.,
; APPLICANT: Spytek, Kimberley A.,
; APPLICANT: Taupier, Jr., Raymond J.,
; APPLICANT: Vernet, Corine A.M.,
; APPLICANT: Zerhusen, Bryan D.,
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-480C
; CURRENT APPLICATION NUMBER: US/10/287,226
; CURRENT FILING DATE: 2002-11-04
; PRIOR APPLICATION NUMBER: 60/334,421
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/354,392
; PRIOR FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: 60/360,148
; PRIOR FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: 60/364,000
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 60/404,821
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: 60/334,526
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/354,409
; PRIOR FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: 60/364,227
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 60/334,027
; PRIOR FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: 60/331,641
; PRIOR FILING DATE: 2001-11-20
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 673
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 669
; LENGTH: 1091
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-287-226-669

Query Match 41.6%; Score 52; DB 15; Length 1091;
Best Local Similarity 40.0%; Pred. No. 24;
Matches 10; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 IVAIAGRLRLMGDPNGELEASEAK 25
Db 910 LVSLTSGLLTIGDRFGGALDAAK 934

RESULT 8
US-10-307-817-86
; Sequence 86, Application US/10307817
; Publication No. US20040058338A1
```



```

; APPLICANT: Vernet, Corine A.M.,
; APPLICANT: Zernusen, Bryan D.,
; APPLICANT: Ziong, Mei
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-480C
; CURRENT APPLICATION NUMBER: US/10/287,226
; CURRENT FILING DATE: 2002-11-04
; PRIOR APPLICATION NUMBER: 60/334,421
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/354,392
; PRIOR FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: 60/360,148
; PRIOR FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: 60/364,000
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 60/404,821
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: 60/334,526
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/354,409
; PRIOR FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: 60/364,227
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 60/334,027
; PRIOR FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: 60/331,641
; PRIOR FILING DATE: 2001-11-20
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 673
; SOFTWARE: CuraSequist version 0.1
; SEQ ID NO 238
; LENGTH: 1095
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-287-226-238

Query Match          41.6%; Score 52; DB 15; Length 1095;
Best Local Similarity 40.0%; Pred. No. 24;
Matches 10; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY      1  IVAIAGRLRMIGDGFNGELESAAK 25
Db      914  LVSLTSGLLTIGDRFGALDAAAK 938

RESULT 11
US-10-287-226-246
; Sequence 246, Application US/10287226
; Publication No. US20040086875A1
; GENERAL INFORMATION:
; APPLICANT: Agee, Michele L.,
; APPLICANT: Alsbrook, John P.,
; APPLICANT: Berghs, Constance,
; APPLICANT: Boldog, Terence,
; APPLICANT: Burgess, Catherine E.,
; APPLICANT: Chant, John S.,
; APPLICANT: Chaudhuri, Amitabha,
; APPLICANT: DiPippo, Vincent A.,
; APPLICANT: Edinger, Shlomit R.,
; APPLICANT: Eisen, Andrew,
; APPLICANT: Ellerman, Karen,
; APPLICANT: Gangolli, Esha A.,
; APPLICANT: Gorman, Linda,
; APPLICANT: Gerlach, Valerie,
; APPLICANT: Ji, Weizhen,
; APPLICANT: Kekuda, Ramesh,
; APPLICANT: Khrantsov, Nikolai,
; APPLICANT: Li, Li,
; APPLICANT: Malyankar, Uriel M.,
; APPLICANT: MacDougall, John R.,
; APPLICANT: Mezes, Peter S.
; APPLICANT: Miller, Charles E.,
; APPLICANT: Millet, Isabelle,

```


This Page Blank (uspio)

Result No.	Score	Query Match	Length	DB	ID	Description
1	52	41.6	1100	2	A35007	ATP citrate (pro-S
2	52	41.6	1105	2	S21173	ATP citrate (pro-S
3	52	41.6	353.9	2	S43048	polyketide synthas
4	49	39.2	579	2	T02574	hypothetical prote
5	47.5	38.0	146	2	H75111	hypothetical prote
6	47	37.6	140	2	C86800	prophage p13 prote
7	47	37.6	140	2	B66683	prophage p11 prote
8	47	37.6	657	2	JC7767	isoamylase (EC 3.1.2
9	46	36.8	360	2	AC2436	polyamine-binding
10	45.5	36.4	148	1	H71021	hypothetical prote
11	45	36.0	132	2	F70557	hypothetical prote
12	45	36.0	297	2	AB0431	lyst-family transc
13	45	36.0	996	2	F86410	protein F3M18.12 (
14	45	36.0	2703	2	H81193	hemagglutinin/hemo
15	44.5	35.6	44	2	B69577	phosphatase (Raps)
16	44.5	35.6	144	4	I51936	hypothetical BCR/A
17	44.5	35.6	308	2	D83858	hypothetical prote
18	44	35.2	285	2	E83599	thiamin biosynthes
19	44	35.2	361	2	AD3198	hypothetical prote
20	44	35.2	412	2	G02453	NN8-4AG - human (f
21	44	35.2	479	2	A64117	serine-type D-Ala-
22	44	35.2	593	2	AC3089	conserved hypothet
23	44	35.2	593	2	E98197	hypothetical prote
24	44	35.2	1099	2	T18713	hypothetical prote
25	44	35.2	1106	2	T29496	hypothetical prote
26	44	35.2	1159	2	S62562	probable nuclear p
27	43.5	34.8	112	2	B24773	protein-tyrosine k
28	43.5	34.8	113	2	C24773	protein-tyrosine k
29	43.5	34.8	137	2	D24773	protein-tyrosine k

A:ATP citrate (pro-S)-lyase - human
 S21173
 C:Species: Homo sapiens (man)
 C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
 R:Accession: S21173
 P:Elshourbagy, N.A.; Near, J.C.; Kmetz, P.J.; Wells, T.N.C.; Groot, P.H.E.;
 Eur. J. Biochem. 204, 491-499, 1992
 A:Title: Cloning and expression of a human ATP-citrate lyase cDNA.
 A:Reference number: S21173; MUID:92174902; PMID:1371749
 A:Accession: S21173
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1105 <ELS>
 A:Cross-references: UNIPROT:P53996; EMBL:X64330; NID:G28934; PIDN:CAA45614.
 C:Superfamily: ATP-citrate synthase

Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 10 RMLGQDFNGELEASAKN 26
: : : : :
7 RIIGDYVNGRLEARIKS 23

Db

RESULT 7
B86683
prophage pil protein 31 [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C;Species: Lactococcus lactis subsp. lactis
C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
C;Accession: B86683
R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
A;Reference number: A86625; MUID:21235186; PMID:11337471
A;Accession: B86683
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-140 <STO>
A;Cross-references: UNIPROT:Q9CIA0; GB:AE005176; PID:gi2723345; PIDN:AAK04564.1; GSPDB:G
A;Experimental source: strain IL1403
C;Genetics:
A;Gene: pil31

Query Match 37.6%; Score 47; DB 2; Length 140;
Best Local Similarity 52.9%; Pred. No. 5.4; Mismatches 3; Indels 0; Gaps 0;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 10 RMLGQDFNGELEASAKN 26
: : : : :
7 RIIGDYVNGRLEARIKS 23

Db

RESULT 8
JC7767
isoamylase (EC 3.2.1.68) - Pectobacterium chrysanthemi
N;Alternate names: glycogen 6-glucanohydrolase
C;Species: Pectobacterium chrysanthemi
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C;Accession: JC7767
R;Lim, W.J.; Park, S.R.; Cho, S.J.; Kim, M.K.; Ryu, S.K.; Hong, S.Y.; Seo, W.T.; Kim, H.
Biochem. Biophys. Res. Commun. 287, 348-354, 2001
A;Title: Cloning and characterization of an intracellular isoamylase gene from Pectobact
A;Reference number: JC7767; PMID:11554733
A;Accession: JC7767
A;Molecule type: DNA
A;Residues: 1-857 <LIM>
A;Cross-references: UNIPROT:Q8KR69; GB:AY044255
A;Experimental source: strain PY35
C;Comment: This enzyme, one of the glycogen-, starch-debranching enzymes, is an intracel
cogen.
C;Genetics:
A;Gene: amyX
C;Superfamily: isoamylase type debranching enzyme
C;Keywords: glycosidase; hydrolase

Query Match 37.6%; Score 47; DB 2; Length 657;
Best Local Similarity 56.2%; Pred. No. 29; Mismatches 3; Indels 0; Gaps 0;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 4 IIAGRLRLMLGDFNGE 19
: : : : :
4 LLAGRPRLGSHFDGE 19

Db

RESULT 9
AC2436
polyamine-binding protein of polyamine ABC transporter all5043 [imported] - Nostoc sp.
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004

C;Accession: AC2436
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anat
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AC2436
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-360 <KUR>
A;Cross-references: UNIPROT:Q8VM93; GB:BA000019; PIDN:BA076742.1; PID:gi7134181; GSPDB:G
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: all5043
C;Superfamily: Escherichia coli spermidine/putrescine-binding protein

Query Match 36.8%; Score 46; DB 2; Length 360;
Best Local Similarity 45.5%; Pred. No. 22; Mismatches 10; Conservative 3; Indels 0; Gaps 0;
Matches 10; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

Qy 4 IIAGRLRLMLGDFNGELEASAK 25
: : : : :
Db 185 VIGAVLRMLGYNSQNEAIK 206

RESULT 10
H71021
hypothetical protein PH1469 - Pyrococcus horikoshii
C;Species: Pyrococcus horikoshii
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 12-Jul-2004
C;Accession: H71021
R;Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekine
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi,
DNA Res. 5, 55-76, 1998
A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a
A;Reference number: AY1000; MUID:98344137; PMID:9679194
A;Accession: H71021
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-148 <KAW>
A;Cross-references: UNIPROT:O59138; GB:AP000006; NID:G3236133; PIDN:BAA30576.1; PID:G325;
A;Experimental source: strain O73
A;Note: this accession replaces an interim accession for a sequence replaced by GenBank
C;Genetics:
A;Gene: PH1469

Query Match 36.4%; Score 45.5; DB 1; Length 148;
Best Local Similarity 40.9%; Pred. No. 9.8; Mismatches 9; Conservative 7; Indels 1; Gaps 1;
Matches 9; Conservative 7; Mismatches 5; Indels 1; Gaps 1;

Qy 1 IVATAGRLRLML-GDQFNGELE 21
: : : : :
Db 5 IVTVKGVKVRWVEGQFIGRIE 26

RESULT 11
F70557
hypothetical protein RV1616 - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C;Accession: F70557
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome s
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: F70557
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-132 <COL>
A;Cross-references: UNIPROT:O06133; GB:Z95554; GB:AL123456; NID:G3261771; PIDN:CAB08893.1

A:Experimental source: strain H37Rv

C:Genetics:

A:Gene: RV1616

Query Match 36.0%; Score 45; DB 2; Length 132;
Best Local Similarity 58.8%; Pred. No. 10;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 7 GRLRLMGDFNGEASAKN 23

Db 55 GGLRMHDLHGELAAAS 71

RESULT 12

AB0431

lysR-family transcription regulatory protein YP03545 [imported] - Yersinia pestis (strain

C:Species: Yersinia pestis

C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004

C:Accession: AB0431

R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.

deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;

il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,

Nature 413, 523-527, 2001

A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A:Reference number: AB0001; MUID:21470413; PMID:11586360

A:Accession: AB0431

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-297 <KUR>

A:Cross-references: UNIPROT:Q8ZB79; GB:AL590842; PIDN:CAC92774.1; PID:G15981467; GSPDB:G

C:Genetics:

A:Gene: YP03545

C:Superfamily: probable transcription regulator ybb5

Query Match 36.0%; Score 45; DB 2; Length 297;
Best Local Similarity 44.4%; Pred. No. 25;
Matches 8; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 4 IIAAGRLMGDFNGELE 21

Db 253 IIAAGRLRVIGPEYSREAD 270

RESULT 13

F86410

protein F3M18.12 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

C:Accession: F86410

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

A:Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: F86410

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-996 <STO>

A:Cross-references: UNIPROT:Q9SGP2; GB:AE005172; NID:G6560764; PIDN:AAF16764.1; GSPDB:GN

C:Genetics:

A:Gene: F3M18.12

A:Map position: 1

C:Superfamily: protein kinase Xa21; leucine-rich alpha-2-glycoprotein repeat homology; P

Query Match 36.0%; Score 45; DB 2; Length 996;
Best Local Similarity 45.0%; Pred. No. 95;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 7 GRLRLMGDFNGEASAKN 26

Db 495 GTLDLHGNOFSGELTSGIKS 514

RESULT 14

H81193

heagglutinin/hemolysin-related protein NME0493 [imported] - Neisseria meningitidis (strain

C:Species: Neisseria meningitidis

C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004

C:Accession: H81193

R:Tettelin, H.; Saunders, N.J.; Heideberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.

Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;

ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.

Science 287, 1809-1815, 2000

A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ven

A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.

A:Reference number: A81000; MUID:20175755; PMID:10710307

A:Accession: H81193

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-2703 <TEP>

A:Cross-references: UNIPROT:Q9K070; GB:AE002405; GB:AE002098; NID:G7225708; PIDN:AAF4092

A:Experimental source: serogroup B, strain MC58

C:Genetics:

A:Gene: NME0493

Query Match 36.0%; Score 45; DB 2; Length 2703;
Best Local Similarity 39.4%; Pred. No. 2.8e+02;
Matches 13; Conservative 5; Mismatches 3; Indels 12; Gaps 2;

QY 6 AGRRLMGDFN-----GELEA-----SAKN 26

Db 552 AAKLRVSGDSFNNTVKGKLAHDLAVNTQTAKN 584

RESULT 15

B69677

phosphatase (RapB) regulator phrE - Bacillus subtilis

C:Species: Bacillus subtilis

C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004

C:Accession: B69677

R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter

C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Choi

A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallerc

iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.;

Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,

A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel,

Y, M.; Ogawa, K.; Ogiwara, A.; Oudga, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle,

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,

A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror,

akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,

T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.

A:Authors: Yoshikawa, H.F.; Zumstein, B.; Yoshikawa, H.; Danchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A:Reference number: A69580; MUID:98044033; PMID:9384377

A:Accession: B69677

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-44 <KUN>

A:Cross-references: UNIPROT:Q32025; GB:Z99117; GB:AL009126; NID:G2634966; PIDN:CAB14525.1

A:Experimental source: strain 168

C:Genetics:

A:Gene: phrE

Query Match 35.6%; Score 44.5; DB 2; Length 44;
Best Local Similarity 33.3%; Pred. No. 3.8;
Matches 8; Conservative 8; Mismatches 7; Indels 1; Gaps 1;

QY 3 AIAAGRLMGDFNGEASAKN 26

Db 11 AVLIG-LAFFGSMYNGEMKEASRN 33

Search completed: November 10, 2004, 14:52:14
Job time : 7.37736 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:38:57 ; Search time 34.634 Seconds
(without alignments)
431.938 Million cell updates/sec

Title: US-10-092-750-44
Perfect score: 125
Sequence: 1 IVAIIGRLRLMLGDFNGEASAKN 26

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	108	86.4	157	2 Q8K186	Q8K186 mus musculus
2	53	42.4	393	2 Q8EAN7	Q8EAN7 shewanella
3	52	41.6	195	2 Q9XK05	Q9XK05 mus musculus
4	52	41.6	592	2 Q9VDN8	Q9VDN8 mus musculus
5	52	41.6	682	2 Q7VZ19	Q7VZ19 bordetella
6	52	41.6	693	2 Q7MDA8	Q7MDA8 bordetella
7	52	41.6	701	2 Q8N9C4	Q8N9C4 homo sapien
8	52	41.6	851	2 Q8V1Q1	Q8V1Q1 rattus norv
9	52	41.6	1091	1 AC1Y MOUSE	Q91V92 mus musculus
10	52	41.6	1100	1 AC1Y RAT	P16638 rattus norv
11	52	41.6	1101	1 AC1Y HUMAN	P53396 homo sapien
12	52	41.6	3519	1 O156 STRAT	Q07017 streptomyce
13	51	40.8	1092	2 Q5DGE7	Q5DGE7 brachydanio
14	49	39.2	248	2 Q700Q0	Q700Q0 pseudomonas
15	49	39.2	248	2 CAF32984	Ca132984 pseudomon
16	49	39.2	278	2 Q7PMH0	Q7PMH0 anopheles g
17	49	39.2	314	2 Q7PJU4	Q7PJU4 anopheles g
18	49	39.2	391	2 Q7JUNC6	Q7JUNC6 drosophila
19	49	39.2	579	2 Q80953	Q80953 arabidopsis
20	49	39.2	607	1 UVRC PSBPK	Q88f17 pseudomonas
21	49	39.2	1086	2 Q7KRA9	Q7KRA9 drosophila
22	49	39.2	1086	2 AAM70940	Aam70940 drosophil
23	49	39.2	1112	2 Q7KN85	Q7KN85 drosophila
24	49	39.2	1112	2 AAD34754	Aad34754 drosophil
25	49	39.2	1118	2 Q7Q0U7	Q7Q0U7 anopheles g
26	49	39.2	3816	2 Q9K1V3	Q9K1V3 streptomyce
27	49	39.2	4150	2 Q9K1V4	Q9K1V4 streptomyce
28	48	38.4	424	2 Q33RK2	Q33RK2 haemophilus
29	48	38.4	723	2 Q9AD52	Q9AD52 streptomyce
30	48	38.4	1054	2 Q8B1X2	Q8B1X2 mus musculus
31	48	38.4	1223	2 Q6P6M5	Q6P6M5 mus musculus

32 48 38.4 1223 2 AAH56933
33 48 38.4 1744 2 Q8CHH7
34 47.5 38.0 146 2 Q9V0U4
35 47.5 38.0 221 2 Q89B89
36 47 37.6 140 2 Q771J0
37 47 37.6 140 2 Q9AYX5
38 47 37.6 140 2 Q9AZ68
39 47 37.6 140 2 Q9AZM9
40 47 37.6 140 2 Q9AZT6
41 47 37.6 140 2 Q9CFR2
42 47 37.6 140 2 Q9CIA0
43 47 37.6 141 2 O53060
44 47 37.6 193 2 O8JGM8
45 47 37.6 292 2 Q7NDC6

ALIGNMENTS

RESULT 1
Q8K186 PRELIMINARY; PRT; 157 AA.
ID Q8K186
AC Q8K186;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
DE Gm566 protein (Fragment).
GN Name=Gm566;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC027668; AAH27668.1; --
FT NONTER 1
SQ SEQUENCE 157 AA; 17078 MW; 92F95902186B5347 CRC64;
Query Match 86.4%; Score 108; DB 2; Length 157;
Best Local Similarity 88.0%; Pred. No. 2e-08;
Matches 22; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 VAIIGRLRLMLGDFNGEASAKN 26

DB 43 VAIIGRLRLMLGDFNGEASAKN 67

```

RESULT 2
Q9EAN7 PRELIMINARY; PRT; 393 AA.
AC Q9EAN7;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Iron-sulfur cluster-binding protein.
GN OrderedLocusNames=SO3861;
OS Shewanella oneidensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Alteromonadaceae; Shewanella.
OX NCBI_TaxID=70863;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MR-1;
RX MEDLINE=22297686; PubMed=12368813; DOI=10.1038/nbt749;
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA Read T.D., Eisen J.A., Seshadri R., Ward N.L., Methe B.A.,
RA Brinkac L.N., Daugherty S.C., DeBoy R.T., Dodson R.J., Durkin A.S.,
RA Hatt D.H., Kolony J.F., Madupu R., Peterson J.D., Umayam L.A.,
RA White O., Wolf A.M., Vamathean J.J., Weidman J.F., Imprim M.,
RA Lee K., Berry K.J., Lee C., Mueller J., Khouri H.M., Gill J.,
RA Uterback T.R., McDonald L.A., Feldblyum T.V., Smith H.O.,
RA Venter J.C., Neilson K.H., Fraser C.M.;
RT "Genome sequence of the dissimilatory metal ion-reducing bacterium
RT Shewanella oneidensis."
RL Nat. Biotechnol. 20:1118-1123 (2002).
DR EMBL; AE015819; AAN56837.1; -.
DR TIGR; SO3861; -.
DR GO; GO:0005489; P:electron transporter activity; IEA.
DR GO; GO:0005506; P:iron ion binding; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001450; 4Fe4S_ferredoxin.
DR InterPro; IPR004453; Fesclus_ferredox.
DR InterPro; IPR009051; Helical_ferredox.
DR Pfam; PF00037; Fer4; 1.
DR TIGRFAMs; TIGR00276; Fesclus binding; 1.
DR PROSITE; PS00198; 4FE4S_FERREDOXIN; 1.
KW 4Fe-4S; Complete proteome; Iron; Iron-sulfur; Metal-binding.
SQ SEQUENCE 393 AA; 43580 MW; 655E08247214B338 CRC64;

Query Match 42.4%; Score 53; DB 2; Length 393;
Best Local Similarity 57.9%; Pred. No. 16;
Matches 11; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 4 ILAGRLMLGDPNGLEA 22
DB 123 LIRARKKLGDQINSIVA 141

RESULT 3
Q99X05 PRELIMINARY; PRT; 195 AA.
AC Q99X05;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Acly protein (Fragment).
GN Name=Acly;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=mix FVB/N;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny K.C., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Krzywinski M.I., Skalska J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=mix FVB/N;
RX TISSUE=Mammary tumor. WAP-TGF alpha model. 7 months old;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC005533; AAH05533.1; -.
DR MGI; MGI:103251; Acly.
DR GO; GO:0004108; P:citrate (S)-synthase activity; IEA.
DR GO; GO:0006099; P:tricarboxylic acid cycle; IEA.
DR InterPro; IPR002020; Citrate_synth.
FT NON_TER 1
SQ SEQUENCE 195 AA; 21693 MW; 68127314F9BB3404 CRC64;

Query Match 41.6%; Score 52; DB 2; Length 195;
Best Local Similarity 40.0%; Pred. No. 12;
Matches 10; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 IVAITAGRLMLGDPNGLEASAK 25
DB 14 LVSSUTSGLLTIGDRFGALDAAK 38.

RESULT 4
Q8VDM8 PRELIMINARY; PRT; 592 AA.
AC Q8VDM8;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Acly protein (Fragment).
GN Name=Acly;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny K.C., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Krzywinski M.I., Skalska J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=mix FVB/N;
RX TISSUE=Mammary tumor. WAP-TGF alpha model. 7 months old;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC005533; AAH05533.1; -.
DR MGI; MGI:103251; Acly.
DR GO; GO:0004108; P:citrate (S)-synthase activity; IEA.
DR GO; GO:0006099; P:tricarboxylic acid cycle; IEA.
DR InterPro; IPR002020; Citrate_synth.
FT NON_TER 1
SQ SEQUENCE 195 AA; 21693 MW; 68127314F9BB3404 CRC64;

```


KW Complete proteome.
SQ SEQUENCE 682 AA; 70475 MW; D8F2FC7842B168B8 CRC64;

Query Match 41.6%; Score 52; DB 2; Length 693;
Best Local Similarity 48.0%; Pred. No. 41;
Matches 12; Conservative 6; Mismatches 5; Indels 2; Gaps 1;

QY 1 IVAITAGRLRMLGDQFN--GELEAS' 23
::: |||::|||::|||::|||
Db 180 LVSLATGGRLVLDGSFSWMGQLGAS 204

RESULT 6
Q7WDAs PRELIMINARY; PRT; 693 AA.

ID Q7WDAs AC Q7WDAs
AC Q7WDAs; AT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative membrane protein.
GN Ordered locus Names=BB3670;
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
ON NCBI_TaxID=518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RB50 / ATCC BAA-588;
RX MEDLINE=22827954; PubMed=12310271; DOI=10.1038/ngl1227;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdeno-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Gobie A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norbertczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Bartell B.G., Maskell D.J.
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640448; CAE35643.1; -
DR InterPro; IPR004814; Oligopept_transpt.
DR InterPro; IPR004813; Telropept_transpt.
DR Pfam; PF03169; OPT_1.
DR TIGRFAMS; TIGR00733; OPT fam: 1.
DR TIGRFAMS; TIGR00728; OPT_sfam; 1.
KW Complete proteome.
SQ SEQUENCE 693 AA; 71698 MW; 71F130E7BD611A88 CRC64;

Query Match 41.6%; Score 52; DB 2; Length 693;
Best Local Similarity 48.0%; Pred. No. 41;
Matches 12; Conservative 6; Mismatches 5; Indels 2; Gaps 1;

QY 1 IVAITAGRLRMLGDQFN--GELEAS 23
::: |||::|||::|||::|||
Db 191 LVSLATGGRLVLDGSFSWMGQLGAS 215

RESULT 7
Q8N9C4 PRELIMINARY; PRT; 701 AA.

ID Q8N9C4 AC Q8N9C4
AC Q8N9C4; AT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein FLJ37765.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
ON NCBI_TaxID=9606;
RN [1]

```

FT  NON_TER      1
SQ  SEQUENCE      1  AA;  92479 MW;  0717514686408CA0  CRC64;

Query Watch      41.6%;  Score 52;  DB 2;  Length 851;
Best Local Similarity 40.0%;  Pred. No. 50;
Matches 10;  Conservative 7;  Mismatches 8;  Indels 0;  Gaps 0;

QY  1  IVAIAGRLRLMLGDDQFNAGELEASAK 25
    : : : : : : : : : : : : : :
Db  670  LVSSLTSGLLITGDRFGGALDAAK 694

RESULT 9
ACLY_MOUSE
ID  ACLY_MOUSE      STANDARD;      PRT;  1091 AA.
AC  Q91V92;
DT  28-FEB-2003 (Rel. 41, Created)
DT  28-FEB-2003 (Rel. 41, Last sequence update)

```

05-JUL-2004 (Rel. 44, Last annotation update)
ATP-citrate synthase (EC 2.3.3.8) (ATP-citrate (pro-S-) -lyase)
(Citrate cleavage enzyme).
Name=Acly;
Mus musculus (Mouse).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A.
STRAIN=ILS, and ISS;
MEDLINE=21363810; PubMed=11471062;
RA Ehringer M.A., Thompson J., Conroy O., Xu Y., Yang F., Canniff J.,
RA Beeson M., Gordon L., Bennett B., Johnson T.E., Sikela J.M.;
RA "High-throughput sequence identification of gene coding variants
RT within alcohol-related QTLs";
RL Mamm. Genome 12:657-663(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22389257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Straubeberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner I., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Helton E., Kettingman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grinwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: ATP citrate-lyase is the primary enzyme responsible for
CC the synthesis of cytosolic acetyl-CoA in many tissues. Has a
CC central role in de novo lipid synthesis. In nervous tissue it may
CC be involved in the biosynthesis of acetylcholine (By similarity).
CC -!- CATALYTIC ACTIVITY: ADP + phosphate + acetyl-CoA + oxaloacetate =
CC ATP + citrate + CoA.
CC -!- SUBUNIT: Homotetramer (By similarity).
CC -!- SIMILARITY: In the N-terminal section; belongs to the
CC succinate/malate CoA ligase beta subunit family.
CC -!- SIMILARITY: In the C-terminal section; belongs to the
CC succinate/malate CoA ligase alpha subunit family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF332052; AAK56081.1; -
CC EMBL; AF332051; AAK56080.1; -
CC EMBL; BC056378; AAK56378.1; -
CC HSRF; P07459; 1SCU.
CC MGD; MGI:103251; Acly.
CC InterPro; IPR002020; Citrate synth.
CC InterPro; IPR003781; CoA binding.
CC InterPro; IPR005811; CoA ligase.
CC InterPro; IPR005810; CoA_lig_alpha.
CC InterPro; IPR005809; CoA_lig_beta.
CC Pfam; PF02629; CoA_Binding; 1.

DR Pfam: PF00549; Ligase CoA; 1.
DR PROSITE; PS01216; SUCINYL_COA_LIG_1; 1.
DR PROSITE; PS00399; SUCINYL_COA_LIG_2; 1.
DR PROSITE; PS01217; SUCINYL_COA_LIG_3; 1.
KW ATP-binding; lipid synthesis; Magnesium; Phosphorylation; Transferase.
FT NP_BIND 691 711 ATP (By similarity).
FT NP_BIND 742 768 ATP (By similarity).
FT DOMAIN 769 779 CoA-binding (Potential).
FT MOD_RES 455 455 Phosphoserine (by PKA) (By similarity).
FT ACT_SITE 750 750 Tele-phosphohistidine intermediate (By similarity).
FT METAL 708 708 Magnesium (By similarity).
SQ SEQUENCE 1091 AA; 119727 MW; 560293D027D797DD CRC64;
Query Match 41.6%; Score 52; DB 1; Length 1091;
Best Local Similarity 40.0%; Pred No; 64;
Matches 10; Conservative 7; Mismatches 8; Indels 0; Gaps 0;
QY 1 IVAVIAGRLRMIGDQFNGLSEASAK 25
Db 910 LVSSLTSGLLTIGDRFGALDAAK 934
RESULT 10
ACLY_RAT STANDARD; PRT; 1100 AA.
ID ACLY_RAT
AC P16638;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE ATP-citrate synthase (EC 2.3.3.8) (ATP-citrate (pro-S-) -lyase)
DE (Citrate cleavage enzyme).
DE Name=Acly;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=90110199; PubMed=2295639;
RA Elshourbagy N.A., Near J.C., Kmetz P.J., Sathe G.M., Southan C.,
RA Strickler J.E., Gross M., Young J.P., Wells T.N.C., Groot P.H.E.;
RT "Rat ATP citrate-lyase. Molecular cloning and sequence analysis of a
RT full-length cDNA and mRNA abundance as a function of diet, organ, and
RT age";
RL J. Biol. Chem. 265:1430-1435(1990).
CC -!- FUNCTION: ATP citrate-lyase is the primary enzyme responsible for
CC the synthesis of cytosolic acetyl-CoA in many tissues. Has a
CC central role in de novo lipid synthesis. In nervous tissue it may
CC be involved in the biosynthesis of acetylcholine.
CC -!- CATALYTIC ACTIVITY: ADP + phosphate + acetyl-CoA + oxaloacetate =
CC ATP + citrate + CoA.
CC -!- ENZYME REGULATION: Major regulation of ATP citrate-lyase activity
CC is probably not by phosphorylation/dephosphorylation but by
CC altering the amount of enzyme.
CC -!- SUBUNIT: Homotetramer.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- PTM: Phosphorylated on two regulatory serines. The ability to the
CC phosphorylate at either one of the regulatory sites depends on the
CC phosphorylation state of the other regulatory site.
CC -!- PTM: The N-terminus is blocked.
CC -!- SIMILARITY: In the N-terminal section; belongs to the
CC succinate/malate CoA ligase beta subunit family.
CC -!- SIMILARITY: In the C-terminal section; belongs to the
CC succinate/malate CoA ligase alpha subunit family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF332052; AAK56081.1; -
CC EMBL; AF332051; AAK56080.1; -
CC EMBL; BC056378; AAK56378.1; -
CC HSRF; P07459; 1SCU.
CC MGD; MGI:103251; Acly.
CC InterPro; IPR002020; Citrate synth.
CC InterPro; IPR003781; CoA binding.
CC InterPro; IPR005811; CoA ligase.
CC InterPro; IPR005810; CoA_lig_alpha.
CC InterPro; IPR005809; CoA_lig_beta.
CC Pfam; PF02629; CoA_Binding; 1.


```

FT CONFLICT      245      245      E -> V (in Ref. 1).
FT CONFLICT      419      423      LGHRP -> WAPA (in Ref. 1).
FT CONFLICT      442      444      SGS -> QRE (in Ref. 1).
FT CONFLICT      457      459      SES -> YESMVDEV (in Ref. 1).
FT CONFLICT      653      656      RPS -> POAA (in Ref. 1).
FT CONFLICT      728      728      C -> S (in Ref. 1).
FT CONFLICT      872      872      V -> A (in Ref. 1).
FT CONFLICT      916      919      AGKD -> TAVE (in Ref. 1).
SQ SEQUENCE      1101 AA; 120825 MW; 5CC770685DCBC23B CRC64;

Query Match          41.6%; Score 52; DB 1; Length 1101;
Best Local Similarity 40.0%; Pred.No. 65;
Matches 10; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY      1 IVAILIAGRLMLGDOFNGELEASAK 25
       :::|::|::|::|::|::|::|
DB     920 LVSSLTSGLLTIIGRPGCALDAAK 944

RESULT 12
OL56_STRAT
ID_OL56_STRAT    STANDARD;      PRT;   3519 AA.
AC      Q07017;
DT      01-NOV-1997 (Rel. 35, Created)
DT      01-NOV-1997 (Rel. 35, Last sequence update)
DT      05-JUL-2004 (Rel. 44, Last annotation update)
DE      Oleandomycin polyketide synthase, modules 5 and 6.
GN      Name-orfB;
OS      Streptomyces antibioticus.
OC      Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC      Streptomycineae; Streptomycetaceae; Streptomycetes.
NCBI_TaxID=1890;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=94150470; PubMed=8107693;
RA      Swan D.G., Rodriguez A.M., Vilches C., Mendez C., Salas J.A.;
RT      Characterisation of a Streptomyces antibiotic gene encoding a type I polyketide synthase which has an unusual coding sequence.";
RT      Mol. Gen. Genet. 242:358-362(1994).
RC      -!- FUNCTION: May be involved in the biosynthesis of the oleandomycin lactone ring.
CC      COFACTOR: Contains 2 covalently bound phosphopantetheines.
CC      -!- SIMILARITY: Contains 2 acyl carrier domains.
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement/ or send an email to license@isb-sib.ch).
-----
CC      EMBL; L09654; AAA19695.1; -.
DR      PIR; S43048; S43048.
DR      HSSP; Q03133; IMO2.
DR      InterPro; IPR039081; ACP_like.
DR      InterPro; IPR001227; Ac_transferase.
DR      InterPro; IPR000794; Ketoacyl_synth.
DR      InterPro; IPR006162; Ppantne_S.
DR      InterPro; IPR006163; Pp_bind.
DR      InterPro; IPR001031; Thioesterase.
DR      Pfam; PF006698; Acyl_transf_1; 2.
DR      Pfam; PF001093; ketoacyl-synt_2.
DR      Pfam; PF02801; ketoacyl-synt_C; 2.
DR      Pfam; PF005550; PP-binding; 2.
DR      Pfam; PF00975; Thioesterase; 1.
DR      PROSITE; PS00075; ACP_DOMAIN; 2.
DR      PROSITE; PS00606; B_KETOACYL_SYNTHASE; 2.
DR      PROSITE; PS00012; PHOSPHOPANTETHEINE; 2.
KW      Acyltransferase; Antibiotic biosynthesis; Multifunctional enzyme;
KW      NADP; Phosphopantetheine; Repeat; Transferase.
FT DOMAIN      1 ? Module 5.
FT DOMAIN      ? 3519 Module 6.

```

Query Match 40.8%; Score 51; DB 2; Length 1092;
Best Local Similarity 36.0%; Pred. No. 92;
Matches 9; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

QY 1 IVAIAGRLRLMGDQFNGLSEASAK 25
DB 911 LISSUTSGLTIGDRFGGALDAAK 935

RESULT 14

QY0000 PRELIMINARY; PRT; 248 AA.
AC Q700Q0;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE UvrC protein (fragment).
GN Name=uvrC;
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=303;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=WCS358;
RA Bertani I., Venturi V.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ629219; CAF32984.1; -
DR InterPro; IPR001943; UvrB/C.
DR InterPro; IPR009055; UvrB C.
DR InterPro; IPR003035; UvrC N.
DR Pfam; PF01541; GIY-YIG; 1.
DR Pfam; PF02151; UVR; 1.
DR SMART; SM00465; GIYC; 1.
DR PROSITE; PS0151; UVR; 1.
DR PROSITE; PS0164; UVR C.1; 1.
FT NON_TER 248
SQ SEQUENCE 248 AA; 27994 MW; 76119C2F2367F398 CRC64;

Query Match 39.2%; Score 49; DB 2; Length 248;
Best Local Similarity 39.1%; Pred. No. 43;
Matches 9; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 2 VAIAGRLRLMGDQFNGLSEASAK 24
DB 195 VNFLEGRSQQLGNELNAEMERAA 217

RESULT 15

CAF32984 PRELIMINARY; PRT; 248 AA.
AC CAF32984;
DT 10-MAR-2004 (TReMBLrel. 27, Created)
DT 10-MAR-2004 (TReMBLrel. 27, Last sequence update)
DT 10-MAR-2004 (TReMBLrel. 27, Last annotation update)
DE UvrC protein (fragment).
GN UVR.
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=303;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=WCS358;
RA Bertani I., Venturi V.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ629219; CAF32984.1; -
FT NON_TER 248

SQ SEQUENCE 248 AA; 27994 MW; 76119C2F2367F398 CRC64;

Query Match 39.2%; Score 49; DB 2; Length 248;
Best Local Similarity 39.1%; Pred. No. 43;
Matches 9; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 2 VAIAGRLRLMGDQFNGLSEASAK 24
DB 195 VNFLEGRSQQLGNELNAEMERAA 217

Search completed: November 10, 2004, 14:49:58
Job time : 36.634 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:44:14 ; Search time 10.6151 Seconds
(without alignments)
181.178 Million cell updates/sec

Title: US-10-092-750-45

Perfect score: 143
Sequence: 1 LALAYSSRQYASALKHAIPIERGIHQH 29

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	51	35.7	224	1	US-08-248-466B-8
2	51	35.7	301	4	US-09-489-039A-11121
3	48.5	33.9	87	3	US-08-134-001C-3005
4	48.5	33.9	224	4	US-09-710-279-3114
5	48.5	33.9	224	4	US-09-710-279-3130
6	48.5	33.9	229	3	US-08-134-001C-5379
7	48.5	33.9	230	3	US-08-134-001C-2907
8	48.5	33.9	230	3	US-08-134-001C-2908
9	48	33.6	463	1	US-07-951-715A-25
10	48	33.6	463	2	US-08-459-448A-25
11	48	33.6	463	3	US-08-459-595A-25
12	48	33.6	463	3	US-08-459-504B-25
13	48	33.6	463	3	US-08-459-444-25
14	48	33.6	463	3	US-08-547-422-25
15	48	33.6	463	4	US-09-988-462-25
16	48	33.6	473	4	US-09-538-092-1353
17	47	32.9	757	4	US-09-494-297A-2
18	46.5	32.5	197	4	US-10-101-464A-491
19	46	32.2	203	4	US-09-270-767-34950
20	46	32.2	203	4	US-09-270-767-50167
21	46	32.2	484	4	US-09-602-787A-26
22	46	32.2	491	3	US-09-362-899-3
23	45	31.5	231	4	US-09-270-767-56909
24	45	31.5	479	4	US-09-270-767-41666
25	45	31.5	623	4	US-09-583-110-4292
26	44.5	31.1	230	3	US-09-134-001C-4549
27	44.5	31.1	276	4	US-09-328-352-5243

28	44.5	31.1	474	4	US-09-248-796A-17867
29	44.5	31.1	2644	4	US-09-029-047C-2
30	44	30.8	86	4	US-09-621-976-7721
31	44	30.8	315	4	US-09-489-039A-8880
32	44	30.8	437	1	US-08-764-343-1
33	44	30.8	437	2	US-08-989-925-3
34	44	30.8	961	4	US-09-252-991A-21535
35	44	30.8	1025	4	US-09-538-092-334
36	43.5	30.4	224	4	US-09-710-279-898
37	43.5	30.4	224	4	US-09-710-279-1134
38	43.5	30.4	243	4	US-09-328-352-7058
39	43	30.1	171	3	US-08-258-287B-40
40	43	30.1	171	3	US-08-368-704C-40
41	43	30.1	240	4	US-09-270-767-36552
42	43	30.1	240	4	US-09-270-767-51769
43	43	30.1	311	2	US-08-602-359A-41
44	43	30.1	355	4	US-09-134-000C-3999
45	43	30.1	495	3	US-08-679-493A-164

ALIGNMENTS

RESULT 1
US-08-248-466B-8
; Sequence 8, Application US/08248466B
; Patent No. 5629182
; GENERAL INFORMATION:
; APPLICANT: CHOPIN, MARIE-CHRISTINE
; APPLICANT: CLUZEL, PIERRE-JEAN
; TITLE OF INVENTION: CLONING OF DNA FRAGMENTS ENCODING A
; TITLE OF INVENTION: MECHANISM FOR RESISTANCE TO BACTERIOPHAGES
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/248,466B
FILING DATE: 24-MAY-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/988,959
FILING DATE: 15-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 90/11381
FILING DATE: 14-SEP-1990
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 5629182man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 384-033-0 PCT FWC CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 224 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein

Query Match 35.7%; Score 51; DB 1; Length 224;

Query Match 33.9%; Score 48.5; DB 3; Length 230;
Best Local Similarity 40.0%; Pred. No. 14;
Matches 12; Conservative 8; Mismatches 7; Indels 3; Gaps 2;

; OTHER INFORMATION: soybean

Query Match 33.6%; Score 48; DB 1; Length 463;
Best Local Similarity 47.4%; Pred. No. 38;

```
Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
;
; OTHER INFORMATION: soybean CDPK as shown in Figure 34."
US-08-459-448A-25
Query Match 33.6%; Score 48; DB 2; Length 463;
Best Local Similarity 47.4%; Pred. No. 38;
Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
QY 5 YSSRQYASALKHIAEIIIE 23
Db 108 HYSEQAARLIKTIVEVVE 126
;
;
RESULT 10
US-08-459-448A-25
; Sequence 25, Application US/08459448A
; Patent No. 5859336
; GENERAL INFORMATION:
; APPLICANT: Koziel, Michael G.
; APPLICANT: Desai, Nalini M.
; APPLICANT: Lewis, Kelly S.
; APPLICANT: Kramer, Vance C.
; APPLICANT: Warren, Gregory W.
; APPLICANT: Evola, Stephen V.
; APPLICANT: Crossland, Lyle D.
; APPLICANT: Rothstein, Steven J.
; APPLICANT: Bowman, Cindy G.
; APPLICANT: Dawson, John L.
; APPLICANT: Dunder, Erik M.
; APPLICANT: Pace, Gary M.
; APPLICANT: Suttie, Janet L.
; TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
; INSECTICIDAL ACTIVITY IN MAIZE
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5859336artis Corporation
; STREET: Patent & Trademark Dept., 520 White Plains
; CITY: Tarrytown
; STATE: New York
; COUNTRY: USA
; ZIP: 10591-9005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,448A
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; FILING DATE: 25-SEP-1992
; APPLICATION NUMBER: US 07/951,715
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/772,027
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Pace, Gary M.
; REGISTRATION NUMBER: 40403
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)541-8689
; TELEFAX: (919)541-8582
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 463 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..463
; OTHER INFORMATION: /note= "protein sequence for
;
;
RESULT 11
US-08-459-595A-25
; Sequence 25, Application US/08459595A
; Patent No. 6018104
; GENERAL INFORMATION:
; APPLICANT: Koziel, Michael G.
; APPLICANT: Desai, Nalini M.
; APPLICANT: Lewis, Kelly S.
; APPLICANT: Kramer, Vance C.
; APPLICANT: Warren, Gregory W.
; APPLICANT: Evola, Stephen V.
; APPLICANT: Crossland, Lyle D.
; APPLICANT: Wright, Martha S.
; APPLICANT: Merlin, Ellis J.
; APPLICANT: Launis, Karen L.
; APPLICANT: Rothstein, Steven J.
; APPLICANT: Bowman, Cindy G.
; APPLICANT: Dawson, John L.
; APPLICANT: Dunder, Erik M.
; APPLICANT: Pace, Gary M.
; APPLICANT: Suttie, Janet L.
; TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
; INSECTICIDAL ACTIVITY IN MAIZE
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6018104artis Corporation
; STREET: Patent & Trademark Dept., 520 White Plains
; CITY: Tarrytown
; STATE: New York
; COUNTRY: USA
; ZIP: 10591-9005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,595A
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/951,715
; FILING DATE: 25-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/772,027
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Pace, Gary M.
; REGISTRATION NUMBER: 40403
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)541-8689
; TELEFAX: (919)541-8582
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 463 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
```

; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..463
; OTHER INFORMATION: /note= "protein sequence for
; OTHER INFORMATION: soybean CDPK as shown in Figure 34."
US-08-459-595A-25

Query Match 33.6%; Score 48; DB 3; Length 463;
Best Local Similarity 47.4%; Pred. No. 38;
Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 5 YYSRQYASALKHIAEII 23
Db 108 HYSERQAARLIKTIIVEVVE 126

RESULT 12
US-08-459-504B-25
; Sequence 25, Application US/08459504B
; Patent No. 6075185
; GENERAL INFORMATION:
; APPLICANT: Koziel, Michael G.
; APPLICANT: Desai, Nalini M.
; APPLICANT: Lewis, Kelly S.
; APPLICANT: Kramer, Vance C.
; APPLICANT: Warren, Gregory W.
; APPLICANT: Evola, Stephen V.
; APPLICANT: Crossland, Lyle D.
; APPLICANT: Wright, Martha S.
; APPLICANT: Merlin, Ellis J.
; APPLICANT: Launis, Karen L.
; APPLICANT: Rothstein, Steven J.
; APPLICANT: Bowman, Cindy G.
; APPLICANT: Dawson, John L.
; APPLICANT: Dunder, Erik M.
; APPLICANT: Pace, Gary M.
; APPLICANT: Suttie, Janet L.
; TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
; TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6075185artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/459,504B
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/459,595
; FILING DATE: 02-JUN-1995
; APPLICATION NUMBER: US 07/951,715
; FILING DATE: 25-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/772,027
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: CGC1577/CIP/DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)541-8587
; TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: 25:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 463 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..463
; OTHER INFORMATION: /note= "protein sequence for
; OTHER INFORMATION: soybean CDPK as shown in Figure 34."
US-08-459-504B-25

Query Match 33.6%; Score 48; DB 3; Length 463;
Best Local Similarity 47.4%; Pred. No. 38;
Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 5 YYSRQYASALKHIAEII 23
Db 108 HYSERQAARLIKTIIVEVVE 126

RESULT 13
US-08-459-444-25
; Sequence 0, Application US/08459444A
; Patent No. 6121014
; GENERAL INFORMATION:
; APPLICANT: Koziel, Michael G.
; APPLICANT: Desai, Nalini M.
; APPLICANT: Lewis, Kelly S.
; APPLICANT: Kramer, Vance C.
; APPLICANT: Warren, Gregory W.
; APPLICANT: Evola, Stephen V.
; APPLICANT: Crossland, Lyle D.
; APPLICANT: Wright, Martha S.
; APPLICANT: Merlin, Ellis J.
; APPLICANT: Launis, Karen L.
; TITLE OF INVENTION: METHOD FOR PRODUCING A PLANT-OPTIMIZED
; TITLE OF INVENTION: NUCLEIC ACID CODING SEQUENCE
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6121014artis Agribusiness Biotechnology Research, Inc.
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/459,444A
; FILING DATE: 02-Jun-1995
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/951,715
; FILING DATE: 25-SEP-1992
; APPLICATION NUMBER: US 07/772,027
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: S-18805/P1/CGC1577/CIP/DIV6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)541-8587
; TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: /note= "protein sequence for
; INFORMATION FOR SEQ ID NO: soybean CDPK as shown in Figure 34."
; SEQUENCE CHARACTERISTICS:
; LENGTH: 463 amino acids

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
FEATURES:
NAME/KEY: Protein
LOCATION: 1..463
SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-08-459-444-25

Query Match 33.6%; Score 48; DB 3; Length 463;
Best Local Similarity 47.4%; Pred. No. 38;
Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 5 YSSRQYASALKHIAEIIIE 23
:|||||:|:|:|:
Db 108 HYSERQAARLKTIVEVVE 126

RESULT 14
US-09-547-422-25
; Sequence 0, Application US/09547422
; Patent No. 6320100
; GENERAL INFORMATION:
; APPLICANT: Koziel, Michael G.
; Desai, Nalini M.
; Lewis, Kelly S.
; Kramer, Vance C.
; Warren, Gregory W.
; Evola, Stephen V.
; Crossland, Lyle D.
; Wright, Martha S.
; Merlin, Ellis J.
; Launis, Karen L.
; TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
; INSECTICIDAL ACTIVITY IN MAIZE
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6320100artis Agribusiness Biotechnology Research, Inc.
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/547,422
; FILING DATE: 11-Apr-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/459,595
; FILING DATE: 02-JUN-1995
; APPLICATION NUMBER: US 07/951,715
; FILING DATE: 25-SEP-1992
; APPLICATION NUMBER: US 07/772,027
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: S-18805H
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)541-8587
; TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: /note= "protein sequence for
; soybean CDPK as shown in Figure 34."
; SEQUENCE CHARACTERISTICS:
; LENGTH: 463 amino acids
; TYPE: amino acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
FEATURES:
NAME/KEY: Protein
LOCATION: 1..463
SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-09-547-422-25

Query Match 33.6%; Score 48; DB 3; Length 463;
Best Local Similarity 47.4%; Pred. No. 38;
Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 5 YSSRQYASALKHIAEIIIE 23
:|||||:|:|:|:
Db 108 HYSERQAARLKTIVEVVE 126

RESULT 15
US-09-988-462-25
; Sequence 25, Application US/09988462
; Patent No. 6720488
; GENERAL INFORMATION:
; APPLICANT: Koziel, Michael G.
; Desai, Nalini M.
; Lewis, Kelly S.
; Kramer, Vance C.
; Warren, Gregory W.
; Evola, Stephen V.
; Crossland, Lyle D.
; Wright, Martha S.
; Merlin, Ellis J.
; Launis, Karen L.
; TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
; INSECTICIDAL ACTIVITY IN MAIZE
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Syngenta Biotechnology, Inc.
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/988,462
; FILING DATE: 20-No. 6720488-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/547,422
; FILING DATE: 11-APR-2000
; APPLICATION NUMBER: US 08/459,504
; FILING DATE: 02-JUN-1995
; APPLICATION NUMBER: US 07/951,715
; FILING DATE: 25-SEP-1992
; APPLICATION NUMBER: US 07/772,027
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: S-18805I
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)541-8587
; TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 463 amino acids
; TYPE: amino acid

```

;
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..463
; OTHER INFORMATION: /note= "protein sequence for
; soybean CDPK as shown in Figure 34."
; SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-09-988-462-25

```

```

Query Match      33.6%; Score 48; DB 4; Length 463;
Best Local Similarity 47.4%; Pred. No. 38;
Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

```

```

Oy      5 YYSSRQYASALKHIAEII 23
       :|||:|:|:|:|:|:|:|:|
Db     108 HYSERQARLIKTIIVEVE 126

```

```

Search completed: November 10, 2004, 14:55:34
Job time : 11.6651 secs

```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2004, 01:28:30 ; Search time 33.5415 Seconds
(without alignments)
305.399 Million cell updates/sec

Title: US-10-092-750-45

Perfect score: 143

Sequence: 1 LALAYSSRQYASALKHIAEIIERGIHQH 29

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	143	100.0	29	14	US-10-092-750-45
2	143	100.0	628	14	Sequence 45, Appl
3	143	100.0	638	14	Sequence 3203, Ap
4	143	100.0	638	17	Sequence 2099, Ap
5	143	100.0	665	15	Sequence 55, Appl
6	54	37.8	131	14	US-10-786-720-55
7	52.5	36.7	109	17	US-10-210-172-74
8	51	35.7	238	16	Sequence 106, App
9	51	35.7	382	17	Sequence 276504
10	51	35.7	398	15	Sequence 32647, A
11	51	35.7	398	15	Sequence 281833
12	51	35.7	398	15	Sequence 53395, A
13	50	35.0	555	15	Sequence 64162, A
					Sequence 72211, A
					Sequence 60515, A

14	34.3	272	14	US-10-369-493-22933	Sequence 22933, A
15	34.3	380	15	US-10-282-122A-49861	Sequence 49861, A
16	34.3	395	15	US-10-424-599-365103	Sequence 265103, A
17	34.3	497	15	US-10-424-599-191824	Sequence 191824, A
18	34.3	502	15	US-10-425-114-40675	Sequence 40675, A
19	34.3	138	15	US-10-425-114-44518	Sequence 44518, A
20	34.3	322	15	US-10-282-122A-53145	Sequence 53145, A
21	34.3	355	16	US-10-437-963-156200	Sequence 156200, A
22	34.3	360	15	US-10-424-599-242079	Sequence 242079, A
23	34.3	379	15	US-10-282-122A-47048	Sequence 47048, A
24	34.3	416	14	US-10-369-493-360	Sequence 260, App
25	34.3	463	10	US-09-988-462-35	Sequence 21338, A
26	34.3	473	14	US-10-320-351-15	Sequence 25, Appl
27	34.3	473	14	US-10-116-275-153	Sequence 15, Appl
28	34.3	508	15	US-10-424-599-242297	Sequence 242297, A
29	34.3	527	15	US-10-425-114-51766	Sequence 51766, A
30	34.3	1101	14	US-10-177-293-106	Sequence 106, App
31	34.3	1101	16	US-10-408-765A-2181	Sequence 2181, App
32	34.3	1360	16	US-10-473-574-12	Sequence 12, Appl
33	34.3	91	17	US-10-425-115-365218	Sequence 365218, A
34	34.3	146	16	US-10-767-701-55611	Sequence 55611, A
35	34.3	169	14	US-10-369-493-10893	Sequence 10893, A
36	34.3	325	14	US-10-369-493-19793	Sequence 19793, A
37	34.3	337	16	US-10-437-963-103961	Sequence 103961, A
38	34.3	549	15	US-10-282-122A-47426	Sequence 47426, A
39	34.3	135	16	US-10-437-963-139454	Sequence 139454, A
40	34.3	167	17	US-10-425-115-281977	Sequence 281977, A
41	34.3	197	14	US-10-101-464A-491	Sequence 491, App
42	34.3	358	16	US-10-322-281-397	Sequence 397, App
43	34.3	382	15	US-10-363-616-269	Sequence 269, App
44	34.3	292	14	US-10-369-493-12712	Sequence 12712, A
45	34.3	46	46		

ALIGNMENTS

RESULT 1

US-10-092-750-45
; Sequence 45, Application US/10092750
; Publication No. US2003032157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; APPLICANT: Wright, Martin C.
; TITLE OF INVENTION: Polypeptides Interactive with BCL-X1
; FILE REFERENCE: 50036/050002 US/10/092,750
; CURRENT APPLICATION NUMBER: US 60/274,526
; PRIOR FILING DATE: 2002-03-07
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 45
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-45

Query Match 100.0%; Score 143; DB 14; Length 29;
Best Local Similarity 100.0%; Pred. No. 8.6e-14;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LALAYSSRQYASALKHIAEIIERGIHQH 29

Db 1 LALAYSSRQYASALKHIAEIIERGIHQH 29

RESULT 2

US-10-094-749-3203
; Sequence 3203, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:

```

; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAWATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3203
; LENGTH: 628
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-3203

Query Match 100.0%; Score 143; DB 14; Length 628;
Best Local Similarity 100.0%; Pred. No. 2.7e-12;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LALAYSSRQYASALKKHIAEIIERGIRQH 29
Db 157 LALAYSSRQYASALKKHIAEIIERGIRQH 185

RESULT 3
US-10-094-749-2099
; Sequence 2099, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAWATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381

Query Match 100.0%; Score 143; DB 14; Length 628;
Best Local Similarity 100.0%; Pred. No. 2.7e-12;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LALAYSSRQYASALKKHIAEIIERGIRQH 29
Db 157 LALAYSSRQYASALKKHIAEIIERGIRQH 185

RESULT 4
US-10-786-720-55
; Sequence 55, Application US/10786720
; Publication No. US2004019181A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: O'Toole, Margot
; APPLICANT: Liu, Wei
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING AND TREATING AUTOIMMUNE
; TITLE OF INVENTION: DISEASES
; FILE REFERENCE: 031896-023000 (AM101331L)
; CURRENT APPLICATION NUMBER: US/10/786,720
; CURRENT FILING DATE: 2004-02-26
; NUMBER OF SEQ ID NOS: 21135
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 55
; LENGTH: 638
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-786-720-55

Query Match 100.0%; Score 143; DB 17; Length 638;
Best Local Similarity 100.0%; Pred. No. 2.7e-12;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LALAYSSRQYASALKKHIAEIIERGIRQH 29
Db 167 LALAYSSRQYASALKKHIAEIIERGIRQH 195

RESULT 5
US-10-210-172-74
; Sequence 74, Application US/10210172
; Publication No. US20040043928A1
; GENERAL INFORMATION:
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Miller, Charles
; APPLICANT: Patturajan, Meera
; APPLICANT: Pena, Carol
; APPLICANT: Rieger, Daniel
; APPLICANT: Shinkets, Richard
; APPLICANT: Zernusen, Bryan
; APPLICANT: Li, Li
; APPLICANT: Ji, Weizhen
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Casman, Stacie
; APPLICANT: Voss, Edward
; APPLICANT: Boldog, Ferenc
; APPLICANT: Gorman, Linda
; APPLICANT: Leite, Mario
; APPLICANT: Vernet, Corine
; APPLICANT: Anderson, David
; APPLICANT: Guo, Xiaojia
; APPLICANT: Zhong, Mei
; APPLICANT: Gerlach, Valerie

```



```

; APPLICANT: Hjalt, Tord
; APPLICANT: Rastelli, Luca
; APPLICANT: Spytek, Kimberly
; APPLICANT: Edinger, Shlomit
; APPLICANT: Ellerman, Karen
; APPLICANT: Maryankar, Griel
; APPLICANT: MacDougall, John
; APPLICANT: Stone, David
; APPLICANT: Alsbrock II, John
; APPLICANT: Lepley, Denise et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-416 A
; CURRENT APPLICATION NUMBER: US/10/210,172
; CURRENT FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: 60/309,501
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 60/323,994
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: 60/373,814
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/310,291
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 60/310,951
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/310,544
; PRIOR FILING DATE: 2001-08-07
; PRIOR APPLICATION NUMBER: 60/311,292
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/311,979
; PRIOR FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: 60/313,201
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 60/312,892
; PRIOR FILING DATE: 2001-08-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 327
; SOFTWARE: Curaseq1ist version 0.1
; SEQ ID NO 74
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-210-172-74

Query Match 100.0%; Score 143; DB 15; Length 665;
Best Local Similarity 100.0%; Pred. No. 2,9e-12;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LALAYSSRQYASALKHIAETIERGIHQ 29
Db 194 LALAYSSRQYASALKHIAETIERGIHQ 222

RESULT 6
US-10-369-493-106
; Sequence 106, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 106
; LENGTH: 131
; TYPE: PRT

; APPLICANT: Aquifex aeolicus
US-10-369-493-106

Query Match 37.8%; Score 54; DB 14; Length 131;
Best Local Similarity 22.4%; Pred. No. 5.5;
Matches 13; Conservative 7; Mismatches 8; Indels 30; Gaps 1;

Qy 1 LALAYSSRQYASALKHIAETIERGIHQ 28
Db 49 LGLEYFKVDYENAIKHLRYELQEDGAAVRTLAKCYVELGEGFQKATEVLEEGIRQ 106

RESULT 7
US-10-425-115-276504
; Sequence 276504, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 276504
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Zea mays
; NAME/KEY: unsure
; FEATURE:
; LOCATION: (1)..(109)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_183754C.1.pep
US-10-425-115-276504

Query Match 36.7%; Score 52.5; DB 17; Length 109;
Best Local Similarity 35.7%; Pred. No. 7.5;
Matches 10; Conservative 7; Mismatches 10; Indels 1; Gaps 1;

Qy 3 LAYSSRQYASALKHIAETIERGIHQ 29
Db 68 IRHHNRQYVFPALSHLSWYVPHGFLRQH 95

RESULT 8
US-10-767-701-32647
; Sequence 32647, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 32647
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C1459_1.pep
US-10-767-701-32647

Query Match 35.7%; Score 51; DB 16; Length 238;
Best Local Similarity 50.0%; Pred. No. 30;
Matches 10; Conservative 3; Mismatches 7; Indels 0; Gaps 0;
```



```
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60515
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Listeria monocytogenes
US-10-282-122A-60515

Query Match          35.0%; Score 50; DB 15; Length 555;
Best Local Similarity 62.5%; Pred. No. 1.1e+02;
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy      11 YASALKHIAEIERGI 26
      ||| ||| ||| ||| |||
Db      175 YASDLSHIAEFGKGV 190

RESULT 14
US-10-369-493-22933
; Sequence 22933, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 22933
; LENGTH: 272

; TYPE: PRT
; ORGANISM: Aeropyrum pernix
US-10-369-493-22933

Query Match          34.3%; Score 49; DB 14; Length 272;
Best Local Similarity 42.3%; Pred. No. 68;
Matches 11; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

Qy      1 LALAYSSRQYASALKHIAEIERGI 26
      : : : : :
Db      24 VGLGTWAIIDYSSALESYVVAIERGI 49

RESULT 15
US-10-282-122A-49861
; Sequence 49861, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 49861
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Burkholderia fungorum
US-10-282-122A-49861

Query Match          34.3%; Score 49; DB 15; Length 380;
Best Local Similarity 63.2%; Pred. No. 99;
Matches 12; Conservative 1; Mismatches 4; Indels 2; Gaps 1;

Qy      9 ROYASALKHIAEIERGI 27
      | | | | |
Db      56 RGYEALKHPEIL--GIR 72

Search completed: November 11, 2004, 02:43:01
```

Fri Nov 12 14:55:29 2004

us-10-092-750-45.rapb

Page 6

Job time : 34.5915 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:40:53 ; Search time 7.11321 Seconds
(without alignments)
392.268 Million cell updates/sec

Title: US-10-092-750-45
Perfect score: 143
Sequence: 1 LALAYSSQYASALKHAEIIRGIHQH 29
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

PIR_79:.*
1: Pirl1.*
2: Pirl2.*
3: Pirl3.*
4: Pirl4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	76	53.1	418	2 F87767	protein F54C1.5 [i
2	54	37.8	131	2 F70422	hypothetical prote
3	53	37.1	368	2 A82249	response regulator
4	52	36.4	1227	2 T49963	hypothetical prote
5	51	35.7	280	2 C64471	hypothetical prote
6	51	35.7	381	2 S37170	repB protein - Lac
7	51	35.7	388	2 S40057	repA protein - Lac
8	51	35.7	403	2 S37168	replication protei
9	51	35.7	550	2 A02852	Na+/Pi-cotransport
10	51	35.7	555	2 B37629	hypothetical prote
11	51	35.7	572	2 H86257	protein F5011.2 [i
12	50.5	35.3	316	2 H90458	hypothetical prote
13	50	35.0	305	2 D69822	ABC transporter (A
14	50	35.0	338	2 B33821	hypothetical prote
15	50	35.0	555	2 A81254	hypothetical prote
16	50	35.0	555	2 A81616	hypothetical prote
17	49	34.3	272	2 C72498	probable stress pr
18	49	34.3	547	2 A95861	hypothetical prote
19	49	34.3	596	2 G75457	tetratricopeptide
20	49	34.3	672	2 A00422	hypothetical prote
21	49	34.3	691	2 S39867	competence protein
22	49	34.3	691	2 G81157	competence protein
23	49	34.3	691	2 B81937	competence protein
24	48.5	33.9	729	2 T41735	probable n-termina
25	48.5	33.9	221	2 B60634	probable transposa
26	48.5	33.9	224	2 C50634	probable transposa
27	48.5	33.9	224	2 A60634	probable transposa
28	48	33.6	145	2 H69051	heterodisulfide re
29	48	33.6	276	2 A80460	conserved hypothe

30	48	33.6	277	2 E72564	hypothetical prote
31	48	33.6	379	1 C70124	conserved hypothe
32	48	33.6	473	1 A53036	Ca2+/calmodulin-de
33	48	33.6	474	1 TVRRC4	Ca2+/calmodulin-de
34	48	33.6	502	2 I52637	calcium-dependent
35	48	33.6	508	1 A43713	hypothetical prote
36	48	33.6	527	2 T03427	hypothetical prote
37	48	33.6	550	2 AF3244	conserved hypothe
38	48	33.6	875	2 F75377	probable helicase
39	48	33.6	2517	2 S58380	probable RNA-direc
40	47.5	33.2	382	2 E85082	hypothetical prote
41	47.5	33.2	382	2 T14186	hypothetical prote
42	47	32.9	206	2 B72236	hypothetical prote
43	47	32.9	301	2 B75153	hypothetical prote
44	47	32.9	368	2 S75132	sensory transducti
45	47	32.9	386	2 JQ1025	replication protei

ALIGNMENTS

RESULT 1

F87767

protein F54C1.5 [imported] - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001

C;Accession: F87767

R;anonymous, The C. elegans Sequencing Consortium.

Science 282, 2012-2018, 1998

A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biology

A;Reference number: A75000; MUID:99069613; PMID:9851916

A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_eleg

A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and

A;Accession: F87767

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-418 <STO>

A;Cross-references: GB:chr_1; PIDN:BA93327.1; PID:g1825574; GSPDB:GN00019; CESP:F54C1.5

C;Genetics:

A;Gene: F54C1.5

A;Map position: 1

Query Match 53.1%; Score 76; DB 2; Length 418;
Best Local Similarity 51.7%; Pred No. 0.0037;
Matches 15; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 LALAYSSQYASALKHAEIIRGIHQH 29

Db 159 IALCHVRGDDYDSALKLISEIIRGVKDH 187

RESULT 2

F70422

hypothetical protein aq_1409 - Aquifex aeolicus

C;Species: Aquifex aeolicus

C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 16-Aug-2004

C;Accession: F70422

R;Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ove

Nature 392, 353-358, 1998

A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.

A;Reference number: A70300; MUID:98196666; PMID:9537320

A;Accession: F70422

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-131 <AQF>

A;Cross-references: UNIPROT:O67408; GB:AE000739; NID:g2983813; PIDN:AA07377.1; PID:g298:

A;Experimental source: strain VFS

C;Genetics:

A;Gene: aq_1409

C;Superfamily: tetratricopeptide repeat homology

F;43-76/Domain: tetratricopeptide repeat homology <TTL>

F;77-110/Domain: tetratricopeptide repeat homology <TT2>

Best Local Similarity 33.3%; Pred. No. 14;
Matches 9; Conservative 11; Mismatches 7; Indels 0; Gaps 0;

QY 3 LAYSSRQYASALKHIAEIIIRGIQH 29
Db 350 LSYVSKQYASAKRNVAKYLKAIQY 376

RESULT 8

S37168
replication protein repB - Lactococcus lactis
C/Species: Lactococcus lactis
C/Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C/Accession: S37168; S44976
R/Seegers, J.; Franke, C.; Venema, G.; Bron, S.; Kiewiet, R.
submitted to the EMBL Data Library, August 1993
A/Description: The majority of lactococcal plasmids belong to one family.
A/Reference number: S37168
A/Accession: S37168
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-403 <SE>
A/Cross-references: UNIPROT:Q48701; EMBL:Z25475
R/Seegers, J.; Franke, C.; Venema, G.; Bron, S.; Kiewiet, R.
submitted to the EMBL Data Library, February 1994
A/Description: The majority of lactococcal plasmids belong to one family.
A/Reference number: S44976
A/Accession: S44976
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 19-403 <SE2>
A/Cross-references: EMBL:Z25475; NID:9452758; PIDN:CAA80964.1; PID:9452759
C/Genetics:
A/Gene: repB
C/Superfamily: Lactococcus lactis replication protein repB

Query Match 35.7%; Score 51; DB 2; Length 403;
Best Local Similarity 33.3%; Pred. No. 15;
Matches 9; Conservative 11; Mismatches 7; Indels 0; Gaps 0;

QY 3 LAYSSRQYASALKHIAEIIIRGIQH 29
Db 368 LSYVSKQYASAKRNVAKYLKAIQY 394

RESULT 9

AD2852
Na+/Pi-cotransporter Atu2245 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C/Species: Agrobacterium tumefaciens
C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C/Accession: AD2852
R/Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, E.W.
A/Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A/Reference number: AB2577; MUID:21608550; PMID:11743193
A/Accession: AD2852
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-550 <KUR>
A/Cross-references: UNIPROT:Q8UD82; GB:AE008688; PIDN:AAL43234.1; PID:gl7740718; GSPDB:G
A/Experimental source: strain C58 (Dupont)
C/Genetics:
A/Gene: Atu2245
A/Map position: circular chromosome

Query Match 35.7%; Score 51; DB 2; Length 550;
Best Local Similarity 62.5%; Pred. No. 20;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 11 YASALKHIAEIIIRGI 26
Db 408 YAINLEHIGDIIIEKGI 423

RESULT 10

B97629
hypothetical protein AGR_C_4085 [imported] - Agrobacterium tumefaciens (strain C58, Cere)
C/Species: Agrobacterium tumefaciens
C/Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C/Accession: B97629
R/Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001
A/Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A/Reference number: A97559; MUID:21608551; PMID:11743194
A/Accession: B97629
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-555 <KUR>
A/Cross-references: UNIPROT:Q8UD82; GB:AE007869; PIDN:AAK87987.1; PID:gl15157399; GSPDB:G
C/Genetics:
A/Gene: AGR_C_4085
A/Map position: circular chromosome

Query Match 35.7%; Score 51; DB 2; Length 555;
Best Local Similarity 62.5%; Pred. No. 21;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 11 YASALKHIAEIIIRGI 26
Db 413 YAINLEHIGDIIIEKGI 428

RESULT 11

H86257
Protein F5011.2 [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cross)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C/Accession: H86257
R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.E.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A/Reference number: A86141; MUID:21616719; PMID:11130712
A/Accession: H86257
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-572 <STO>
A/Cross-references: UNIPROT:Q9LNB6; GB:AE005172; NID:gs8778620; PIDN:AAF79628.1; GSPDB:GN
C/Genetics:
A/Gene: F5011.2
A/Map position: 1

Query Match 35.7%; Score 51; DB 2; Length 572;
Best Local Similarity 32.0%; Pred. No. 21;
Matches 8; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 5 YVSSRQYASALKHIAEIIIRGIQH 29
Db 393 FFKQKYPEAKHYTEAKRNPNH 417

RESULT 12

H90458
hypothetical protein SSO2816 [imported] - Sulfolobus solfataricus
C/Species: Sulfolobus solfataricus

C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
C;Accession: H90458
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aweyaz, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, H.
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A;Description: Sulfolobus solfataricus complete genome.
A;Reference number: A99139
A;Accession: H90458
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-316 <KUP>
A;Cross-references: UNIPROT:Q97V20; GB:AE006641; NID:gl3816166; PIDN:AAK42927.1; GSPDB:G
C;Genetics:
A;Gene: SSO2816

Query Match 35.3%; Score 50.5; DB 2; Length 316;
Best Local Similarity 40.6%; Pred. No. 14;
Matches 13; Conservative 7; Mismatches 7; Indels 5; Gaps 1;
QY 1 LALAYYS-----SRQYASALKHIAEIIERGIR 27
DB 164 TALAISGGYTFVARGAYDVKHLKEIIKKAIK 195

RESULT 13
D69822
ABC transporter (ATP-binding protein) homolog yhcH - Bacillus subtilis.
C;Species: Bacillus subtilis
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 16-Aug-2004
C;Accession: D69822
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 330, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningstein, G.; Krog, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauele
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron
akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpatra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A;Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A69580; MUID:198044033; PMID:9384377

A;Accession: D69822
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-305 <KUN>
A;Cross-references: UNIPROT:P54592; GB:Z99108; GB:AL009126; NID:G2633055; PIDN:CAB12736.
A;Experimental source: strain 168
C;Genetics:
A;Gene: yhcH
C;Superfamily: ATP-binding cassette homology
C;Keywords: ATP; nucleotide binding; P-loop
F:20-209/Domain: ATP-binding cassette homology <ABC>
F:37-44/Region: nucleotide-binding motif A (P-loop)

Query Match 35.0%; Score 50; DB 2; Length 305;
Best Local Similarity 50.0%; Pred. No. 15;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 8 SRQYASALKHIAEIIIE 23
DB 69 TREYAKAIKHGAIVE 84

RESULT 14
B83861
hypothetical protein BHI690 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans

C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C;Accession: B83861
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hiran
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: B83861
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-338 <STO>
A;Cross-references: UNIPROT:Q9KC84; GB:AP001512; GB:BA000004; NID:gl0174030; PIDN:BA8054
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BHI690

Query Match 35.0%; Score 50; DB 2; Length 338;
Best Local Similarity 40.0%; Pred. No. 17;
Matches 10; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 1 LALAYSSQYASALKHIAEIIERG 25
DB 181 LGMAIFLQHYALAIPCFMEVIKDG 205

RESULT 15
AB1254
hypothetical protein lmo1434 [imported] - Listeria monocytogenes (strain EGD-e)
C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C;Accession: AB1254
R;Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker,
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.;
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkut, G.; Madueno, E.; Maitournam, A.; Mat
ok, C.; Schluteter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AB1254
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-555 <GLA>
A;Cross-references: UNIPROT:Q8Y767; GB:NC_003210; PIDN:CAC9512.1; PID:gl6410863; GSPDB:C
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: lmo1434
C;Superfamily: conserved hypothetical protein MG139

Query Match 35.0%; Score 50; DB 2; Length 555;
Best Local Similarity 62.5%; Pred. No. 29;
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 11 YASALKHIAEIIERGII 26
DB 175 YASDLSHIAEFGKGV 190

Search completed: November 10, 2004, 14:52:15
Job time : 8.11321 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:38:57 ; Search time 38.6302 Seconds
(without alignments)
431.938 Million cell updates/sec

Title: US-10-092-750-45
Perfect score: 143
Sequence: 1 LALAYSSRQYASALKHAEIIRGIQH 29

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt 02.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	143	100.0	638	Q96NE6	Q96ne6 homo sapien
2	143	100.0	665	Q81VP2	Q81vp2 homo sapien
3	143	100.0	665	Q8N4P2	Q8n4p2 homo sapien
4	143	100.0	678	Q86WT1	Q86wt1 homo sapien
5	135	94.4	664	Q8C0X1	Q8c0x1 mus musculus
6	135	94.4	664	Q99J38	Q99j38 m riken cdn
7	135	94.4	664	Q9CY00	Q9cy00 mus musculus
8	118	82.5	651	Q61NU8	Q61nu8 xenopus lae
9	118	82.5	651	A8H72174	A8h72174 xenopus l
10	81	56.6	634	Q7PRA4	Q7pra4 anopheles g
11	78	54.5	652	Q9VK41	Q9vk41 drosophila
12	76	53.1	597	Q817G3	Q817g3 caenorhabdi
13	76	53.1	656	Q817G4	Q817g4 caenorhabdi
14	55	38.5	516	Q9WNB1	Q9wnb1 agrobacteri
15	54	37.8	131	Q67408	Q67408 apuifex aeo
16	54	37.8	209	Q8BP35	Q8bp35 xanthomonas
17	54	37.8	209	Q8PNA0	Q8pna0 xanthomonas
18	54	37.8	663	Q8NM61	Q8nm61 dictyosteli
19	54	37.8	1279	Q7QVW8	Q7qv8 giardia lam
20	53	37.1	368	Q7MIU6	Q7miu6 vibrio vuln
21	53	37.1	368	Q8DB27	Q8db27 vibrio vuln
22	53	37.1	368	Q8KT55	Q8kt55 vibrio chol
23	52	36.4	247	Q8PSM9	Q8psm3 methanocarc
24	52	36.4	365	Q7N0T3	Q7n0t3 photorhabdu
25	52	36.4	495	1 THSB_SULAC	Q9v2t4 sulfolobus
26	52	36.4	595	Q8T020	Q8t020 drosophila
27	52	36.4	615	Q9W3Y6	Q9w3y6 drosophila
28	52	36.4	615	2 AAF46175	Aaf46175 drosophil
29	52	36.4	1227	2 Q9LXFP4	Q9lxf4 arabidopsis
30	51.5	36.0	101	2 Q8DA67	Q8da67 mus musculus
31	51.5	36.0	226	2 Q8CID2	Q8cid2 mus musculus

32	51.5	36.0	295	2	Q9AJL9	Q9ajl9 hydrogenoba
33	51	35.7	174	2	Q7N0J4	Q7n0j4 photorhabdu
34	51	35.7	252	2	Q8Y3F6	Q8y3f6 ralstonia s
35	51	35.7	280	1	UPFS_METJA	Q58767 methanococc
36	51	35.7	375	2	Q6BXK8	Q6bxk8 debaryomyce
37	51	35.7	381	2	Q48703	Q48703 lactococcus
38	51	35.7	383	2	Q845D7	Q845d7 lactococcus
39	51	35.7	385	2	P94889	P94889 lactococcus
40	51	35.7	385	2	Q48701	Q48701 lactococcus
41	51	35.7	385	2	Q9QE4	Q9qe4 lactococcus
42	51	35.7	385	2	Q9AIQ4	Q9aiq4 lactococcus
43	51	35.7	388	2	Q48681	Q48681 lactococcus
44	51	35.7	389	2	Q48698	Q48698 lactococcus
45	51	35.7	550	2	Q8UD82	Q8ud82 agrobacteri

ALIGNMENTS

RESULT 1
Q96NE6 PRELIMINARY; PRT; 638 AA.
AC Q96NE6
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Hypothetical protein FLJ30990.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RC SEQUENCE FROM N.A.
RP TISSUE=Lung;
RX PubMed=14702039;
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahari K.,
RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
RA Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura M.,
RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,
RA Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
RA Tanai H., Kimata M., Watanabe M., Hiraka S., Chiba Y., Ishida S.,
RA Ono Y., Takiguchi S., Watanabe S., Yoshida M., Hotuta T., Kusano J.,
RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,
RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
RA Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA Moriya S., Snoch A., Mizoguchi H., Ichihara T., Shiohata N., Sano S.,
RA Nakagawa S., Snoch A., Mizoguchi H., Ichihara T., Shiohata N., Sano S.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
cDNAs".
RL Nat. Genet. 36:40-45 (2004).
DR EMBL; AB055552; BAB70953.1; -.
DR InterPro; IPR001440; TPR.
DR InterPro; IPR008941; TPR-like.
DR Pfam; PF00515; TPR; 3.
DR SMART; SM00028; TPR; 3.
DR PROSITE; PS50293; TPR_REGION; 1.
KW Repeat; TPR repeat.

```

SQ SEQUENCE 638 AA; 73146 MW; 16AB964E71ACFP893 CRC64;
Query Match 100.0%; Score 143; DB 2; Length 638;
Best Local Similarity 100.0%; Pred. No. 6.3e-12;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LALAYSSRQYASALKHIAEIIERGIQH 29
DB 167 LALAYSSRQYASALKHIAEIIERGIQH 195

RESULT 2
Q81VP2 PRELIMINARY; PRT; 665 AA.
AC Q81VP2;
DT 01-VAR-2003 (TReMBLrel. 23, Created)
DT 01-VAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-VAR-2004 (TReMBLrel. 26, Last annotation update)
DE Hypothetical protein FLJ13946.
GN Name=FLJ13946;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RN "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[2]
SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX Submitted (JAN-2003) to the EMBL/GenBank/DBSJ databases.
RA EMBL; BC042848; AAH42848.1; -.
RA InterPro; IPR008940; Pfam; PF00515; TPR-like.
RA InterPro; IPR001440; TPR.
RA InterPro; IPR008941; TPR-like.
RA Pfam; PF00515; TPR; 3.
RA SMART; SM00028; TPR; 4.
RA PROSITE; PS0293; TPR_REGION; 1.
KW Hypothetical protein; Repeat; TPR repeat.
SQ SEQUENCE 665 AA; 76177 MW; DC44E17689642854 CRC64;

Query Match 100.0%; Score 143; DB 2; Length 665;
Best Local Similarity 100.0%; Pred. No. 6.3e-12;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LALAYSSRQYASALKHIAEIIERGIQH 29
DB 194 LALAYSSRQYASALKHIAEIIERGIQH 222

RESULT 3
Q81VP2 PRELIMINARY; PRT; 665 AA.
AC Q81VP2;
DT 01-VAR-2003 (TReMBLrel. 23, Created)
DT 01-VAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-VAR-2004 (TReMBLrel. 26, Last annotation update)
DE Hypothetical protein FLJ13946.
GN Name=FLJ13946;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RN "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[2]
SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX Submitted (JAN-2003) to the EMBL/GenBank/DBSJ databases.
RA EMBL; BC042848; AAH42848.1; -.
RA InterPro; IPR008940; Pfam; PF00515; TPR-like.
RA InterPro; IPR001440; TPR.
RA InterPro; IPR008941; TPR-like.
RA Pfam; PF00515; TPR; 3.
RA SMART; SM00028; TPR; 4.
RA PROSITE; PS0293; TPR_REGION; 1.
KW Hypothetical protein; Repeat; TPR repeat.
SQ SEQUENCE 665 AA; 76177 MW; DC44E17689642854 CRC64;

Query Match 100.0%; Score 143; DB 2; Length 665;
Best Local Similarity 100.0%; Pred. No. 6.3e-12;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LALAYSSRQYASALKHIAEIIERGIQH 29
DB 194 LALAYSSRQYASALKHIAEIIERGIQH 222

RESULT 4
Q86WT1 PRELIMINARY; PRT; 678 AA.
AC Q86WT1;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE FLJ13946 protein (Fragment).
GN Name=FLJ13946;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RN "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[2]
SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX Submitted (JUL-2002) to the EMBL/GenBank/DBSJ databases.
RA EMBL; BC033795; AAH33795.1; -.
RA InterPro; IPR008940; Pfam; PF00515; TPR-like.
RA InterPro; IPR001440; TPR.
RA InterPro; IPR008941; TPR-like.
RA Pfam; PF00515; TPR; 5.
RA SMART; SM00028; TPR; 4.
RA PROSITE; PS0293; TPR_REGION; 2.
KW Repeat; TPR repeat.
SQ SEQUENCE 665 AA; 76079 MW; 651E5DE826A3E070 CRC64;

Query Match 100.0%; Score 143; DB 2; Length 665;
Best Local Similarity 100.0%; Pred. No. 6.3e-12;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LALAYSSRQYASALKHIAEIIERGIQH 29
DB 194 LALAYSSRQYASALKHIAEIIERGIQH 222

RESULT 5
Q86WT1 PRELIMINARY; PRT; 678 AA.
AC Q86WT1;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE FLJ13946 protein (Fragment).
GN Name=FLJ13946;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RN "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[2]
SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX Submitted (JUL-2002) to the EMBL/GenBank/DBSJ databases.
RA EMBL; BC033795; AAH33795.1; -.
RA InterPro; IPR008940; Pfam; PF00515; TPR-like.
RA InterPro; IPR001440; TPR.
RA InterPro; IPR008941; TPR-like.
RA Pfam; PF00515; TPR; 5.
RA SMART; SM00028; TPR; 4.
RA PROSITE; PS0293; TPR_REGION; 2.
KW Repeat; TPR repeat.
SQ SEQUENCE 665 AA; 76079 MW; 651E5DE826A3E070 CRC64;

Query Match 100.0%; Score 143; DB 2; Length 665;
Best Local Similarity 100.0%; Pred. No. 6.3e-12;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LALAYSSRQYASALKHIAEIIERGIQH 29
DB 194 LALAYSSRQYASALKHIAEIIERGIQH 222

```

```

RX MEDLINE=21085660; PubMed=11217851;
RT RIKEN FANTOM Consortium;
RA "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
[3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RA The FANTOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of
  60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
[4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RA MEDLINE=20499374; PubMed=11042159;
RX Carninci P., Shibata Y., Hayashizaki Y.;
RA Konno H., Okazaki Y., Muramatsu M., Sugahara Y., Shibata K., Itoh M.,
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
  prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
[5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RA MEDLINE=20530913; PubMed=11076861;
RX Shibata K., Itoh M., Aizawa K., Nagaoaka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuina T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishire T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwara S., Inoue K., Togawa Y., Iwata M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kita A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
  sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
[6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saichoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toyota T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK029535; BAC26502.1; -
DR MGD; MGI:1926052; 493050613Rik.
DR InterPro; IPR008940; Ptenyl_trans.
DR InterPro; IPR001440; TPR.
DR InterPro; IPR008941; TPR-like.
DR Pfam; PF00515; TPR; 4.
DR SMART; SM00028; TPR; 4.
DR PROSITE; PS00293; TPR_REGION; 1.
KW Hypothetical protein; Repeat; TPR repeat.
SQ SEQUENCE 664 AA; 76221 MW; 70FA38E48D95F3A5 CRC64;

Query Match 94.4%; Score 135; DB 2; Length 664;
Best Local Similarity 93.1%; Pred.No. 9.2e-11;
Matches 27; Conservative 1; Mismatches 1; Indels 0; Gaps

QY 1 LALAYSSRQYASAKGHIAETIERGIQH 29
Db 193 LALAYSSRQYAPALGHIAETIERGIQH 221

RESULT 6
Q99J38 PRELIMINARY; PRT; 664 AA.
ID Q99J38

```

PRT; 664 AA.

AC Q99J38;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE RIKEN CDNA 4930506L13 (Mus musculus 9.5 days embryo parthenogenote
 DE cDNA, RIKEN full-length enriched library, clone: B13032P08
 DE product: hypothetical Tetrairicopeptide repeat (TPR) structure
 DE containing protein, full insert sequence)
 GN Name=4930506L13Rik;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NMRI; TISSUE=Mammary tumor. WAP-Tag model. 5 months old;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Klausner R.L., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Shat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marsina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Diatchenko M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettner M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield V.S.,
 RA Krzywinski M.I., Skalska U., Smalls D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NMRI; TISSUE=Mammary tumor. WAP-Tag model. 5 months old;
 RA Strausberg R.;
 RL Submitted (MAR-2001) to the ENBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Parthenogenote;
 RX MEDLINE=99279253; PubMed=10349636;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning.";
 RL Meth. Enzymol. 303:19-44 (1999).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Parthenogenote;
 RX MEDLINE=21085660; PubMed=11217851;
 RA RIKEN FANTOM Consortium;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690 (2001).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Parthenogenote;
 RA The FANTOM Consortium;
 RA The RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:553-573 (2002).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Parthenogenote;
 RX MEDLINE=20499374; PubMed=11042159;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630 (2000).

RN [7]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Parthenogenote;
 RX MEDLINE=20530913; PubMed=11076861;
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
 RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771 (2000).
 RN [8]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Parthenogenote;
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
 RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
 RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
 RA Tagawa A., Takahashi F., Takaku-Akaiura S., Takeda Y., Tanaka T.,
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
 RL Submitted (JUL-2001) to the ENBL/GenBank/DBJ databases.
 DR EMBL; BC004773; AA04773.1; -;
 DR EMBL; AK045096; BAC3220.1; -;
 DR MGD; MGI:1926052; 4930506L13Rik.
 DR InterPro; IPR008940; Prenyl_trans.
 DR InterPro; IPR001440; TPR.
 DR Pfam; PF00515; TPR; 4.
 DR SMART; SM00028; TPR; 4.
 DR PROSITE; P850293; TPR_REGION; 1.
 KW Hypothetical protein; Repeat; TPR repeat.
 SQ SEQUENCE 664 AA; 76237 MW; F9C3C7863231CEA2 CRC64;
 Query Match 94.4%; Score 135; DB 2; Length 664;
 Best Local Similarity 93.1%; Pred. No. 9, 2e-11;
 Matches 27; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 LALAYSSQYASALKHIAEIIERGIQHQ 29
 DB 193 LALAYSSQYAPALKHIAEIIERGIQHQ 221
 RESULT 7
 Q9CY00 PRELIMINARY; PRT; 664 AA.
 AC Q9CY00;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Mus musculus 13 days embryo liver cDNA, RIKEN full-length enriched
 DE library, clone: 2510042P03 product: 4930506L13Rik PROTEIN HOMOLOG.
 GN Name=2510042P03Rik;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Liver;
 RX MEDLINE=99279253; PubMed=10349636;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning.";
 RL Meth. Enzymol. 303:19-44 (1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Liver;
 RN [3]

RX MEDLINE=21085660; PubMed=11217851;
 RA RIKEN FANTOM Consortium;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Liver;
 RX MEDLINE=20499374; PubMed=11042159;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630(2000).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Liver;
 RX MEDLINE=20530913; PubMed=11076661;
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
 RA Konno H., Akiyama J., Nishi K., Katsunai T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watabiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa Y., Tanaka T., Matsura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771(2000).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Liver;
 RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
 RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
 RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
 RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
 RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
 RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
 RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
 RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
 RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
 RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
 RA Muramatsu M., Hayashizaki Y.;
 RT Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AK011097; BAB27394.1; -
 DR MGD; MGI:1919671; 2510042P03Rik.
 DR InterPro; IPR001440; TPR.
 DR InterPro; IPR008941; TPR-like.
 DR Pfam; PF00515; TPR; 3.
 DR SMART; SM00028; TPR; 4.
 DR PROSITE; PS0293; TPR_REGION; 1.
 DR Repeat; TPR repeat.
 SQ SEQUENCE 664 AA; 76095 MW; 8BCB2D1F1A7099F3 CRC64;

Query Match 94.4%; Score 135; DB 2; Length 664;
 Best Local Similarity 93.1%; Pred. No. 9.2e-11;
 Matches 27; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LALAYSSFOYASALKHIAIIRGIHQH 29
 |||||
 DB 193 LALAYSSFOYAPALKHADIIRGIHQH 221
 |||||

RESULT 8
 Q6INU8 PRELIMINARY; PRT; 651 AA.
 ID Q6INU8
 AC Q6INU8;

DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE MGC80272 protein.
 GN Name=MGC80272;
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Ovary;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hong F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.J., Skalek U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Ovary;
 RX MEDLINE=22341132; PubMed=12454917;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.;
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 RT initiative.";
 RL Dev. Dyn. 225:384-391(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Ovary;
 RX Klein S., Strausberg R.;
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC072174; AAH72174.1; -
 DR InterPro; IPR001440; TPR.
 DR InterPro; IPR008941; TPR-like.
 DR Pfam; PF00515; TPR; 2.
 DR SMART; SM00028; TPR; 4.
 DR PROSITE; PS00005; TPR; 1.
 DR PROSITE; PS0293; TPR_REGION; 2.
 DR Repeat; TPR repeat.
 SQ SEQUENCE 651 AA; 75558 MW; AF0608F6F7986590 CRC64;

Query Match 82.5%; Score 118; DB 2; Length 651;
 Best Local Similarity 75.9%; Pred. No. 2.6e-08;
 Matches 22; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 LALAYSSROYASALKHIAIIRGIHQH 29
 |||||
 DB 180 IALCYYSMKQYAPALKHADIIRGIHQH 208
 |||||

RESULT 9
 AAH72174 PRELIMINARY; PRT; 651 AA.
 ID AAH72174
 AC AAH72174;
 DT 01-JUN-2004 (TrEMBLrel. 27, Created)
 DT 01-JUN-2004 (TrEMBLrel. 27, Last sequence update)

```

DE Hypothetical protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=22341132; PubMed=12454517;
RA Klein S.L., Strausberg R.L., Wagner L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usslin T.B., Toshlyyiki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RA Klein S., Strausberg R.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC072174; AAH72174.1;
KW Hypothetical protein.
SQ SEQUENCE 651 AA; 75558 MW; AF060876F7986590 CRC64;

Query Match 82.5%; Score 118; DB 2; Length 651;
Best Local Similarity 75.9%; Pred. No. 2.6e-08;
Matches 22; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 LALAYSSQVYASALKHIAIEIIRGIROH 29
Db :|||||:|||||:|||||:|||||:|||||:
180 IALCYVMQYAPALKHIAIEIIRGIRESH 208

RESULT 10
Q7PRA4 PRELIMINARY; PRT; 634 AA.
AC Q7PRA4;
DT 01-MAR-2004 (TEMBLrel. 26, Created)
DT 01-MAR-2004 (TEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE ENSANGP0000024665 (Fragment).
GN Name=ENSANGG0000020924.
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB0100859; EAA07585.2; -.
DR InterPro; IPR001440; TPR.
DR PROSITE; PS0005; TPR; 1.
DR PROSITE; PS0293; TPR REGION; 1.
FT NON_TER 1
FT NON_TER 634
FT NON_TER 634
SQ SEQUENCE 634 AA; 69956 MW; 7954B31DDE29C4DA CRC64;

Query Match 56.8%; Score 81; DB 2; Length 634;
Best Local Similarity 53.6%; Pred. No. 0.006;
Matches 15; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY 2 ALAYSSQVYASALKHIAIEIIRGIROH 29
Db :|||||:|||||:|||||:|||||:|||||
167 ALCHYRKENSQALYIAIEIRGINH 194

RESULT 11
Q9VK41 PRELIMINARY; PRT; 652 AA.
ID Q9VK41
AC Q9VK41;
DT 01-MAY-2000 (TEMBLrel. 13, Created)
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE CGS142-PA.
GN ORFNames=CGS142;
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Aghayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Murthy D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodgett, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
```


RT investigating biology. The C. elegans Sequencing Consortium.":
RL Science 282:2012-2018(1998).
RN [2]

RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Miller N., Bradshaw H., Wansley P.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
RN [3]

RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterson R.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
RN [4]

RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Wilson R.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
RN [5]

RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Wilson R.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
RN [6]

DR EMBL; U88165; AA012420.1; -;
DR WormPep; F54C1.5a; CB33108.
DR InterPro; IPR001440; TPR.
DR SMART; SM0008941; TPR-like.
DR InterPro; IPR008941; TPR.
DR SMART; SM00028; TPR; 3
DR PROSITE; PS00293; TPR_REGION; 1.
KW Hypothetical protein.

SQ SEQUENCE 656 AA; 75607 MW; 6C10AACD65DC3C20 CRC64;

Query Match 53.1%; Score 76; DB 2; Length 656;
Best Local Similarity 51.7%; Pred. No. 0.033;
Matches 15; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 LALAYSSRQYASALKHAEIIRGIQ 29

Db 181 IALCHVRGDDYDSALKLSEIINRGVXDH 209

RESULT 14

Q9WMB1 PRELIMINARY; PRT; 516 AA.

AC C9WMB1;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Agrobacterium tumefaciens.
OG Plasmid pTiChry5.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=358;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Chry5;
RA Oger P.M., Farrand S.K.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF065246; AAD31607.1; -;
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0015321; F:sodium-dependent phosphate transporter acti. .; IEA.
DR GO; GO:0006817; P:phosphate transport; IEA.
DR InterPro; IPR003841; Na/Pi_cotransp.
DR Pfam; PF02690; Na_Pi_cotrans; 1.
DR TIGRFAM; TIGR01013; 2a58; 1.
KW Hypothetical protein; Plasmid.
FT NON TER 1
SQ SEQUENCE 516 AA; 55197 MW; 563AF8A98BAEFB8A CRC64;

Query Match 38.5%; Score 55; DB 2; Length 516;
Best Local Similarity 55.8%; Pred. No. 29;
Matches 10; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 11 YASALKHAEIIRGIQ 28
Db 374 YATNLEHIGDIIEKGLLQ 391

RESULT 15

O67408 PRELIMINARY; PRT; 131 AA.

AC O67408;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein aq.1409.
GN Ordered locus names=AQ_1409;
OS Aquifex aeolicus.
OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE-98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Aujay M., Huber R.,
RA Feldman R.A., Short J.M., Olsen G.J., Swanson R.V.;
RL "The complete genome of the hyperthermophilic bacterium Aquifex
aeolicus.";
RT Nature 392:353-358(1998).
DR EMBL; AE000739; AAC07377.1; -;
DR PIR; F70422; F70422; TPR.
DR InterPro; IPR001440; TPR.
DR Pfam; PF00515; TPR; 2.
DR PROSITE; PS00005; TPR; 1.
DR PROSITE; PS00293; TPR_REGION; 1.
KW Complete proteome; Hypothetical protein.

SQ SEQUENCE 131 AA; 15774 MW; 49B428E0D8580709 CRC64;

Query Match 37.8%; Score 54; DB 2; Length 131;
Best Local Similarity 22.4%; Pred. No. 10;
Matches 13; Conservative 7; Mismatches 8; Indels 30; Gaps 1;

QY 1 LALAYSSRQYASALKHAEIIRGIQ 28

Db 49 LGLEYFKVDYENAIKHLERYLELQDEGAAYRTLAKCYEELGEFQKAEVLEEGIRQ 105

Search completed: November 10, 2004, 14:50:00
Job time : 40.6302 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:44:14 ; Search time 13.9094 Seconds
(without alignments)
181.178 Million cell updates/sec

Title: US-10-092-750-46

Perfect score: 210

Sequence: 1 AAMLLDRGTCDLWINEMSLHKIVQDVYGTTPPHS 38

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/iaa/5A.COMB.pdp:*
- 2: /cgn2_6/ptodata/1/iaa/5B.COMB.pdp:*
- 3: /cgn2_6/ptodata/1/iaa/6A.COMB.pdp:*
- 4: /cgn2_6/ptodata/1/iaa/6B.COMB.pdp:*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS.COMB.pdp:*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pdp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57.5	27.4	1124	4	US-09-270-767-46552
2	56	26.7	334	4	US-09-543-681A-4753
3	52	24.8	219	4	US-09-540-236-3745
4	52	24.8	399	3	US-08-765-907A-10
5	52	24.8	1037	4	US-09-340-620A-55
6	51.5	24.5	261	4	US-09-377-285B-44
7	51	24.3	298	4	US-09-248-796A-14868
8	51	24.3	566	3	US-09-028-934-31
9	51	24.3	567	1	US-08-258-261B-4
10	51	24.3	567	1	US-08-456-837-4
11	51	24.3	567	1	US-08-457-342-4
12	51	24.3	567	1	US-08-457-646A-4
13	51	24.3	567	1	US-08-458-076A-4
14	51	24.3	567	1	US-08-457-335A-4
15	51	24.3	567	2	US-08-729-214-4
16	51	24.3	567	3	US-09-028-934-4
17	50.5	24.0	263	3	US-09-134-001C-4638
18	50.5	24.0	303	4	US-09-543-681A-6212
19	50.5	24.0	1431	4	US-09-538-092-1198
20	50	23.8	178	4	US-09-489-039A-10005
21	50	23.8	649	4	US-09-418-963-2
22	49	23.3	138	4	US-09-270-767-41247
23	49	23.3	138	4	US-09-270-767-56463
24	49	23.3	152	4	US-09-621-976-7078
25	49	23.3	197	4	US-09-328-352-4665
26	49	23.3	517	4	US-09-257-835B-21
27	49	23.3	2037	4	US-09-543-681A-5538

28	49	23.3	2414	1	US-08-227-536-2	Sequence 2, Appli
29	49	23.3	2414	4	US-09-538-092-1269	Sequence 1289, Ap
30	49	23.3	2414	5	PCT-US95-04682-2	Sequence 2, Appli
31	49	23.3	2441	1	US-08-194-468-2	Sequence 2, Appli
32	49	23.3	2441	3	US-08-961-739-2	Sequence 8, Appli
33	49	23.3	2441	3	US-09-514-247A-8	Sequence 2, Appli
34	49	23.3	2441	4	US-09-686-316-2	Sequence 10, Appli
35	49	23.3	2442	3	US-09-514-247A-10	Sequence 1370, Ap
36	49	23.3	2442	4	US-09-538-092-1370	Sequence 9445, Ap
37	48	22.9	425	4	US-09-489-039A-9445	Sequence 20, Appl
38	48	22.9	520	4	US-09-457-825B-20	Sequence 35, Appl
39	48	22.9	565	3	US-09-028-934-35	Sequence 2, Appli
40	48	22.9	1288	1	US-07-727-814B-2	Sequence 2, Appli
41	48	22.9	1288	1	US-08-258-614-2	Sequence 64, Appli
42	48	22.9	1911	4	US-09-854-856-64	Sequence 48, Appli
43	48	22.9	1939	4	US-09-854-856-48	Sequence 12, Appli
44	48	22.9	1971	4	US-09-854-856-32	Sequence 16, Appli
45	48	22.9	1999	4	US-09-854-856-16	

ALIGNMENTS

RESULT 1

US-09-270-767-46552

; Sequence 46552, Application US/09270767

; Patent No. 6703491

; GENERAL INFORMATION:

; APPLICANT: Homburger et al.

; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

; FILE REFERENCE: File Reference: 7326-094

; CURRENT APPLICATION NUMBER: US/09/270,767

; CURRENT FILING DATE: 1999-03-17

; NUMBER OF SEQ ID NOS: 62517

; SOFTWARE: Patentin Ver. 2.0

; SEQ ID NO 46552

; LENGTH: 1124

; TYPE: PRT

; ORGANISM: Drosophila melanogaster

; FEATURE:

; OTHER INFORMATION: Xaa means any amino acid

US-09-270-767-46552

Query Match 27.4%; Score 57.5; DB 4; Length 1124;
Best Local Similarity 35.3%; Pred. No. 20;
Matches 12; Conservative 5; Mismatches 6; Indels 11; Gaps 2;

Cy 5 LDRRGTECDLWINEMSLHKIVQDVYGTTPPHS 38

Db 569 IDQSSRL-CWLS-----DTVTTPPHS 591

RESULT 2

US-09-543-681A-4753

; Sequence 4753, Application US/09543681A

; Patent No. 6605709

; GENERAL INFORMATION:

; APPLICANT: GARY BRETON

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS

; FILE REFERENCE: 2709.1002-001

; CURRENT APPLICATION NUMBER: US/09/543,681A

; CURRENT FILING DATE: 2000-04-05

; PRIOR APPLICATION NUMBER: US 60/128,706

; PRIOR FILING DATE: 1999-04-09

; NUMBER OF SEQ ID NOS: 8344

; SEQ ID NO 4753

; LENGTH: 334

; TYPE: PRT

; ORGANISM: Proteus mirabilis

US-09-543-681A-4753

Query Match 26.7%; Score 56; DB 4; Length 334;

; ORGANISM: Rattus norvegicus
 US-09-377-285B-44
 Query Match 24.5%; Score 51.5; DB 4; Length 261;
 Best Local Similarity 32.4%; Pred. No. 25;
 Matches 11; Conservative 9; Mismatches 13; Indels 1; Gaps 1;

QY 5 LDRRGTECDLWIN-EMSLHLKIVQDVYGTTPHPH 37
Db 103 LRRGGEERMLVDEEYEQDLRDMEARPPPH 136

RESULT 7
US-09-248-796A-14868
; Sequence 14868, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 14868
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-14868

Query Match 24.3%; Score 51; DB 4; Length 298;
Best Local Similarity 28.6%; Pred. No. 36;
Matches 12; Conservative 5; Mismatches 11; Indels 14; Gaps 2;

QY 9 GTECDL-----W--INEMSLHLKIVQDVYGTTPHPH 36
Db 74 GTESSIQHHYAHNSLFWKLFNQSLHLKIVSKNKNPSHPH 115

RESULT 8
US-09-028-934-31
; Sequence 31, Application US/09028934
; Patent No. 6117670
; GENERAL INFORMATION:
; APPLICANT: Ligon, James M.
; APPLICANT: Hill, Dwight S.
; APPLICANT: Lam, Steven T.
; APPLICANT: Hammer, Philip B.
; APPLICANT: van Bee, Karl-Heinz
; APPLICANT: Kirner, Sabine
; APPLICANT: Young, Thomas R.
; TITLE OF INVENTION: Pyroclitrin Biosynthesis Genes and Uses
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6117670artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/028,934
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/729,214
; FILING DATE: 09-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/258,261

QY 16 INEMSLHLKIVQDVYGTTPHPH 37
Db 58 IPETSLMNRITADRYGIPELDH 79

RESULT 9
US-08-258-261B-4
; Sequence 4, Application US/08258261B
; Patent No. 5639949
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas M.
; APPLICANT: Ligon, James M.
; APPLICANT: Beck, James Joseph
; APPLICANT: Hill, Dwight Steven
; APPLICANT: Ryals, John Andrew
; APPLICANT: Gaffney, Thomas Deane
; APPLICANT: Lam, Stephen Ting
; APPLICANT: Hammer, Phillip B.
; APPLICANT: Uknes, Scott Joseph
; TITLE OF INVENTION: Genes for the synthesis of
; antipathogenic substances
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ciba-Geigy Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/258,261B
; FILING DATE: 08-JUN-1994
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/457,205
; FILING DATE: 01-JUN-1995
; ATORNEY/AGENT INFORMATION:
; NAME: Elmer, James Scott
; REGISTRATION NUMBER: 36,129
; REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8614
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 567 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

Query Match 24.3%; Score 51; DB 3; Length 566;
Best Local Similarity 45.5%; Pred. No. 74;
Matches 10; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 16 INEMSLHLKIVQDVYGTTPHPH 37
Db 58 IPETSLMNRITADRYGIPELDH 79

US-09-028-934-31
; FILING DATE: 08-JUN-1994
; ATORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: CGC1506/CIP7
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 566 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-028-934-31

MOLECULE TYPE: protein
US-08-258-261B-4
Query Match 24.3%; Score 51; DB 1; Length 567;
Best Local Similarity 45.5%; Pred. No. 74;
Matches 10; Conservative 4; Mismatches 8; Indels 0; Gaps 0;
QY 16 INEMSLHKKIVQDVYGTTPPH 37
DB 58 IPETSLMNRRIADRYGIPDLH 79

RESULT 10
US-08-456-837-4
Sequence 4, Application US/08456837
Patent No. 5643774
GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Ligon, James M.
APPLICANT: Beck, James Joseph
APPLICANT: Hill, Dwight Steven
APPLICANT: Ryals, John Andrew
APPLICANT: Gafney, Thomas Deane
APPLICANT: Lam, Stephen Ting
APPLICANT: Hammer, Phillip E.
APPLICANT: Uknes, Scott Joseph
TITLE OF INVENTION: Genes for the synthesis of
antipathogenic substances
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ciba-Geigy Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: NY
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/456,837
FILING DATE: 01-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/457,205
FILING DATE: 01-JUN-1995
APPLICATION NUMBER: 08/258,261
FILING DATE: 08-Jun-1994
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8614
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 567 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-456-837-4

Query Match 24.3%; Score 51; DB 1; Length 567;
Best Local Similarity 45.5%; Pred. No. 74;
Matches 10; Conservative 4; Mismatches 8; Indels 0; Gaps 0;
QY 16 INEMSLHKKIVQDVYGTTPPH 37
DB 58 IPETSLMNRRIADRYGIPDLH 79

RESULT 11
US-08-457-342-4
Sequence 4, Application US/08457342
Patent No. 5662898
GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Ligon, James M.
APPLICANT: Beck, James Joseph
APPLICANT: Hill, Dwight Steven
APPLICANT: Ryals, John Andrew
APPLICANT: Gafney, Thomas Deane
APPLICANT: Lam, Stephen Ting
APPLICANT: Hammer, Phillip E.
APPLICANT: Uknes, Scott Joseph
TITLE OF INVENTION: Genes for the synthesis of
antipathogenic substances
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ciba-Geigy Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: NY
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/457,342
FILING DATE: 01-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/457,205
FILING DATE: 01-JUN-1995
APPLICATION NUMBER: 08/258,261
FILING DATE: 08-Jun-1994
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8614
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 567 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-457-342-4

Query Match 24.3%; Score 51; DB 1; Length 567;
Best Local Similarity 45.5%; Pred. No. 74;
Matches 10; Conservative 4; Mismatches 8; Indels 0; Gaps 0;
QY 16 INEMSLHKKIVQDVYGTTPPH 37
DB 58 IPETSLMNRRIADRYGIPDLH 79

RESULT 12
US-08-457-646A-4
Sequence 4, Application US/08457646A
Patent No. 5679560
GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Ligon, James M.
APPLICANT: Beck, James Joseph
APPLICANT: Hill, Dwight Steven
APPLICANT: Ryals, John Andrew

APPLICANT: Gaffney, Thomas Deane
APPLICANT: Lam, Stephen Ting
APPLICANT: Hammer, Phillip E.
APPLICANT: Uknes, Scott Joseph
TITLE OF INVENTION: Genes for the synthesis of
antipathogenic substances
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ciba-Geigy Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: NY
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/457,646A
FILING DATE: 01-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/457,205
FILING DATE: 01-JUN-1995
APPLICATION NUMBER: 08/258,261
FILING DATE: 08-Jun-1994
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
TELEPHONE: 919-541-8614
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 567 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-457-646A-4

Query Match 24.3%; Score 51; DB 1; Length 567;
Best Local Similarity 45.5%; Pred. No. 74;
Matches 10; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

OY 16 INEMSLHKIVDVGTGTHPPH 37
DB 58 IPETSLMNRRIADRYGIPDLH 79

RESULT 13
US-08-458-076A-4
Sequence 4, Application US/08/458076A
Patent No. 5698425
GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Ligon, James M.
APPLICANT: Beck, James Joseph
APPLICANT: Hill, Dwight Steven
APPLICANT: Ryals, John Andrew
APPLICANT: Gaffney, Thomas Deane
APPLICANT: Lam, Stephen Ting
APPLICANT: Hammer, Phillip E.
APPLICANT: Uknes, Scott Joseph
TITLE OF INVENTION: Genes for the synthesis of
antipathogenic substances
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ciba-Geigy Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne

STATE: NY
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,076A
FILING DATE: 01-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/457,205
FILING DATE: 01-JUN-1995
APPLICATION NUMBER: 08/258,261
FILING DATE: 08-Jun-1994
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
TELEPHONE: 919-541-8614
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 567 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-458-076A-4

Query Match 24.3%; Score 51; DB 1; Length 567;
Best Local Similarity 45.5%; Pred. No. 74;
Matches 10; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

OY 16 INEMSLHKIVDVGTGTHPPH 37
DB 58 IPETSLMNRRIADRYGIPDLH 79

RESULT 14
US-08-457-335A-4
Sequence 4, Application US/08/457335A
Patent No. 5723759
GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Ligon, James M.
APPLICANT: Beck, James Joseph
APPLICANT: Hill, Dwight Steven
APPLICANT: Ryals, John Andrew
APPLICANT: Gaffney, Thomas Deane
APPLICANT: Lam, Stephen Ting
APPLICANT: Hammer, Phillip E.
APPLICANT: Uknes, Scott Joseph
TITLE OF INVENTION: Genes for the synthesis of
antipathogenic substances
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ciba-Geigy Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: NY
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/457,335A
FILING DATE: 01-JUN-1995

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2004, 01:28:30 ; Search time 43.9509 Seconds
(without alignments)
305.399 Million cell updates/sec

Title: US-10-092-750-46

Perfect score: 210

Sequence: 1 AAMLLDRRGTECDLWLNEMSLHKIVQDVYGTPTPHPS 38

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09D_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	210	100.0	38	14	US-10-092-750-46
2	58	27.6	1011	14	US-10-156-761-8999
3	57	27.1	184	14	US-10-080-170-608
4	57	27.1	184	16	US-10-080-170-608
5	57	27.1	184	17	US-10-468-356-608
6	54.5	26.0	102	16	US-10-437-963-127540
7	54.5	26.0	1755	16	US-10-437-963-188840
8	54	25.7	210	17	US-10-425-115-313042
9	54	25.7	275	14	US-10-106-698-5169
10	54	25.7	476	14	US-10-104-047-2206
11	54	25.7	484	15	US-10-282-122A-74677
12	54	25.7	517	14	US-10-094-749-2115
13	53	25.2	303	17	US-10-425-115-333859

14	53	25.2	378	16	US-10-437-963-159960	Sequence 159960, Ap
15	53	25.2	468	15	US-10-335-977-8588	Sequence 8588, Ap
16	53	25.2	522	15	US-10-335-977-8589	Sequence 8589, Ap
17	53	25.2	742	14	US-10-156-761-12153	Sequence 12153, A
18	53	25.2	1121	16	US-10-437-963-159959	Sequence 159959, A
19	52.5	25.0	168	15	US-10-424-539-176444	Sequence 176444, A
20	52.5	25.0	1439	16	US-10-437-963-138825	Sequence 138825, A
21	52	24.8	269	16	US-10-369-493-195253	Sequence 195253, A
22	52	24.8	523	14	US-10-437-963-10395	Sequence 10395, A
23	52	24.8	1037	9	US-09-728-721-55	Sequence 55, Appl
24	52	24.8	1037	14	US-10-295-981-55	Sequence 55, Appl
25	51.5	24.5	129	17	US-10-425-115-264527	Sequence 264527, A
26	51.5	24.5	261	14	US-10-192-381-44	Sequence 44, Appl
27	51.5	24.5	382	14	US-10-369-493-14266	Sequence 14266, A
28	51	24.3	64	17	US-10-425-115-268841	Sequence 268841, A
29	51	24.3	135	9	US-09-783-436-12	Sequence 12, Appl
30	51	24.3	486	15	US-10-282-122A-72523	Sequence 72523, A
31	51	24.0	567	14	US-10-148-907A-5	Sequence 5, Appl
32	50.5	24.0	181	15	US-10-424-539-158052	Sequence 158052, A
33	50.5	24.0	190	14	US-10-050-882-57	Sequence 57, Appl
34	50.5	24.0	311	15	US-10-282-122A-67409	Sequence 67409, A
35	50.5	24.0	336	14	US-10-080-170-25	Sequence 25, Appl
36	50.5	24.0	336	16	US-10-080-170-25	Sequence 25, Appl
37	50.5	24.0	336	17	US-10-468-356-25	Sequence 25, Appl
38	50.5	24.0	509	14	US-10-369-493-11768	Sequence 11768, A
39	50.5	24.0	837	14	US-10-369-493-18586	Sequence 18586, A
40	50.5	24.0	1431	14	US-10-352-684A-38	Sequence 38, Appl
41	50.5	24.0	1431	14	US-10-045-400C-5	Sequence 5, Appl
42	50.5	24.0	1431	14	US-10-295-027-1235	Sequence 1235, Ap
43	50	23.8	93	16	US-10-437-963-157588	Sequence 157588, A
44	50	23.8	165	17	US-10-425-115-332283	Sequence 332283, A
45	50	23.8	637	15	US-10-425-114-57951	Sequence 57951, A

ALIGNMENTS

RESULT 1

US-10-092-750-46
; Sequence 46, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; APPLICANT: Wright, Martin C.
; TITLE OF INVENTION: Polypeptides Interactive with BCL-X1
; FILE REFERENCE: 50036/05002
; CURRENT APPLICATION NUMBER: US/10/092,750
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/274,526
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-46

Query Match 100.0%; Score 210; DB 14; Length 38;
Best Local Similarity 100.0%; Pred. No. 2.4e-21;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAMLLDRRGTECDLWLNEMSLHKIVQDVYGTPTPHPS 38
DB 1 AAMLLDRRGTECDLWLNEMSLHKIVQDVYGTPTPHPS 38

RESULT 2

US-10-156-761-8999
; Sequence 8999, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:

```

; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 8999
; LENGTH: 1011
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-8999

```

```

Query Match 27.6%; Score 58; DB 14; Length 1011;
Best Local Similarity 37.1%; Pred. No. 61;
Matches 13; Conservative 8; Mismatches 12; Indels 2; Gaps 1;

```

```

QY 1 AAMLLDRRGTECDLWINEMSLHKKIVQDVYGTGPH 35
DB 916 AMLLDQGNSEAM--TEVNAHRIYTDKYSRHP 948

```

```

RESULT 3
US-10-080-170-608
; Sequence 608, Application US/10080170
; Publication No. US20030129601A1
; GENERAL INFORMATION:
; APPLICANT: COLE, S.T.
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
; TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; FILE REFERENCE: 03495.0218
; CURRENT APPLICATION NUMBER: US/10/080,170
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/270,123
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 652
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 608
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-080-170-608

```

```

Query Match 27.1%; Score 57; DB 14; Length 184;
Best Local Similarity 35.3%; Pred. No. 13;
Matches 12; Conservative 7; Mismatches 15; Indels 0; Gaps 0;

```

```

QY 2 AMLDRRGTECDLWINEMSLHKKIVQDVYGTGPH 35
DB 122 AFILDRMSQQVDADHRVALLRKTVGETWGLPSP 155

```

```

RESULT 4
US-10-080-170-608
; Sequence 608, Application US/10080170
; Publication No. US20040121322A9
; GENERAL INFORMATION:
; APPLICANT: COLE, S.T.
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
; TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; FILE REFERENCE: 03495.0218
; CURRENT APPLICATION NUMBER: US/10/080,170

```

```

; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/270,123
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 652
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 608
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-080-170-608

```

```

Query Match 27.1%; Score 57; DB 16; Length 184;
Best Local Similarity 35.3%; Pred. No. 13;
Matches 12; Conservative 7; Mismatches 15; Indels 0; Gaps 0;

```

```

QY 2 AMLDRRGTECDLWINEMSLHKKIVQDVYGTGPH 35
DB 122 AFILDRMSQQVDADHRVALLRKTVGETWGLPSP 155

```

```

RESULT 5
US-10-468-356-608
; Sequence 608, Application US/10468356
; Publication No. US20040197896A1
; GENERAL INFORMATION:
; APPLICANT: COLE, STEWART
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
; TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; FILE REFERENCE: 05394.0019
; CURRENT APPLICATION NUMBER: US/10/468,356
; CURRENT FILING DATE: 2003-08-19
; PRIOR APPLICATION NUMBER: 10/080,170
; PRIOR FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: 60/270,123
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 655
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 608
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-468-356-608

```

```

Query Match 27.1%; Score 57; DB 17; Length 184;
Best Local Similarity 35.3%; Pred. No. 13;
Matches 12; Conservative 7; Mismatches 15; Indels 0; Gaps 0;

```

```

QY 2 AMLDRRGTECDLWINEMSLHKKIVQDVYGTGPH 35
DB 122 AFILDRMSQQVDADHRVALLRKTVGETWGLPSP 155

```

```

RESULT 6
US-10-437-963-127540
; Sequence 127540, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Bing
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 127540

```



```
; LENGTH: 102
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(102)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_29982C.1.pep
US-10-437-963-127540

Query Match      26.0%; Score 54.5; DB 16; Length 102;
Best Local Similarity 24.3%; Pred. No. 15;
Matches 9; Conservative 12; Mismatches 15; Indels 1; Gaps 1;

Qy  3 MLLDRRGTECDLWINEMS-LLHKIVQDVYGTTPPHPS 38
Db  14 LILDRASNGDVYTHQNIQIESAETLYGNMHPKH 50

RESULT 7
US-10-437-963-188840
; Sequence 188840, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 188840
; LENGTH: 1755
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_85405C.1.pep
US-10-437-963-188840

Query Match      26.0%; Score 54.5; DB 16; Length 1755;
Best Local Similarity 38.5%; Pred. No. 3,3e+02;
Matches 10; Conservative 7; Mismatches 8; Indels 1; Gaps 1;

Qy  14 LWI-NEMSLHKIVQDVYGTTPPHPS 38
Db  1070 IWGNNTALQKLVNHLATPLGHS 1095

RESULT 8
US-10-425-115-313042
; Sequence 313042, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 313042
; LENGTH: 210
```

```
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_48552C.1.pep
US-10-425-115-313042

Query Match      25.7%; Score 54; DB 17; Length 210;
Best Local Similarity 52.9%; Pred. No. 39;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy  22 LHKIVQDVYGTTPPHPS 38
Db  1 MHAYILDFFYTCHPPHS 17

RESULT 9
US-10-106-698-5169
; Sequence 5169, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 5169
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (154)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (181)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (184)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (265)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-5169

Query Match      25.7%; Score 54; DB 14; Length 275;
Best Local Similarity 34.1%; Pred. No. 52;
Matches 14; Conservative 9; Mismatches 12; Indels 6; Gaps 3;

Qy  3 MLLDRRGTECDL---WINEM--SLLHKIVQDVYGTTPPHPS 38
Db  3 LVLERRSGDRDLEPDWLAQLRQLQKQVAGDI-GDPHPTRS 42

RESULT 10
US-10-104-047-2206
; Sequence 2206, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1el full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
```

[illegible]

```
Query Match      25.2%; Score 53; DB 17; Length 303;
Best Local Similarity 37.9%; Pred. No. 79;
Matches 11; Conservative 6; Mismatches 8; Indels 4; Gaps 2;

Qy 12 CDLWINEMS-LLHKIVQD---VYGTTPHP 36
Db 5 CDLWLTGLGCLIFSLVLRQPLGCLYGTSDPP 33

RESULT 14
US-10-437-963-159960
; Sequence 159960, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 159960
; LENGTH: 378
; TYPE: PET
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_59285C.1.pep
US-10-437-963-159960

Query Match      25.2%; Score 53; DB 16; Length 378;
Best Local Similarity 45.5%; Pred. No. 1e+02;
Matches 10; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

Qy 8 RGTECDLWINEMSLHLKIVQDV 29
Db 107 RSSFCCLWLNCFNSHNKIVQOI 128

RESULT 15
US-10-335-977-8588
; Sequence 8588, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 06/993,002
```

```
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 8588:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 468 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3) LOCATION 1...468
; SEQUENCE DESCRIPTION: SEQ ID NO: 8588:
US-10-335-977-8588

Query Match      25.2%; Score 53; DB 15; Length 468;
Best Local Similarity 54.2%; Pred. No. 1.3e+02;
Matches 13; Conservative 3; Mismatches 6; Indels 2; Gaps 1;

Qy 2 AMLDRLRG--TECDLWINEMSLHLH 23
Db 285 AMLFDRNGVETECDLKVDCKELLN 308

Search completed: November 11, 2004, 02:43:02
Job time : 45.0009 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:40:53 ; Search time 9.32076 Seconds
(without alignments)
392.268 Million cell updates/sec

Title: US-10-092-750-46

Perfect score: 210

Sequence: 1 AAMLLDRRGTECDLWINEMSLHKIVQDVYGTTPPHPS 38

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79:*

1: Piri:*

2: Pir2:*

3: Pir3:*

4: Pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57.5	27.4	308	2 D83858	hypothetical prote
2	57.5	27.4	574	2 T40213	hypothetical prote
3	57	27.1	184	2 G70956	hypothetical prote
4	57	27.1	1116	2 B70476	hypothetical prote
5	55	26.2	312	2 T20932	hypothetical prote
6	54	25.7	245	2 A99163	dipeptide transpor
7	54	25.7	245	2 A93124	hypothetical prote
8	53	25.2	522	2 B71978	hypothetical prote
9	52	24.8	131	2 AF3338	hypothetical prote
10	51.5	24.5	297	2 T08589	hypothetical prote
11	51	24.3	186	2 A81272	probable dCNP deam
12	51	24.3	259	2 A98302	oligopeptide trans
13	51	24.3	259	2 AC2981	hypothetical prote
14	50.5	24.0	181	2 G85930	interacts with sec
15	50.5	24.0	181	2 E91085	Syd protein [impor
16	50.5	24.0	181	2 A59944	syd protein - Esch
17	50.5	24.0	336	2 C85921	conserved hypothet
18	50.5	24.0	561	2 AE2649	alpha-glucosidase
19	50.5	24.0	594	2 D74331	probable alpha-glu
20	50.5	24.0	837	2 H84239	dimethylaldehyde
21	50	23.8	190	1 D64566	dCNP deaminase (EC
22	50	23.8	227	2 C84431	hypothetical prote
23	50	23.8	232	2 D71157	probable haloacid
24	50	23.8	300	2 T00274	hypothetical prote
25	50	23.8	361	2 H96634	hypothetical prote
26	50	23.8	459	2 AC0075	probable membrane
27	50	23.8	483	2 S75369	hypothetical prote
28	50	23.8	579	2 AE1855	hypothetical prote
29	50	23.8	821	2 T16412	hypothetical prote

30 50 23.8 916 2 T04752
31 49 23.3 182 2 AE1957
32 49 23.3 223 2 T16335
33 49 23.3 283 2 F72476
34 49 23.3 427 2 F83984
35 49 23.3 433 2 F96974
36 49 23.3 2414 2 A54277
37 49 23.3 2440 2 S39162
38 49 23.3 2441 2 S39161
39 48.5 23.1 80 2 B85567
40 48.5 23.1 379 2 F95153
41 48.5 23.1 429 2 G87636
42 48 22.9 165 2 E70366
43 48 22.9 274 2 A84936
44 48 22.9 277 2 I51258
45 48 22.9 332 2 T35917

ALIGNMENTS

RESULT 1

D83858
hypothetical protein BHL68 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C;Accession: D83858
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hii
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: D83858
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-308 <STO>
A;Cross-references: UNIPROT:Q9KCA5; GB:AP001512; GB:BA000004; NID:g10174030; PIDN:BA005
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BHL68

Query Match 27.4%; Score 57.5; DB 2; Length 308;
Best Local Similarity 46.2%; Pred. No. 3.7;
Matches 12; Conservative 5; Mismatches 4; Indels 5; Gaps 2;

Cy 13 DLWINEMSLHKIVQDVYGTTPPHPS 38
|||: : ||| :
Db 189 DLWVS-LCEAHRI----YGTTHPPNS 209

RESULT 2

T40213
hypothetical protein SPBC31F10.10c - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T40213
R;Wood, V.; Rajandream, M.A.; Barrell, B.G.; Pohl, T.
submitted to the EMBL Data Library, August 1997
A;Reference number: Z21913
A;Accession: T40213
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-574 <WOO>
A;Cross-references: UNIPROT:P87311; EMBL:Z97204; PIDN:CAB10087.1; GSPDB:GN00067; SPDB:S
A;Experimental source: strain 972h-; cosmid c31F10
C;Genetics:
A;Gene: SPDB:SPBC31F10.10c
A;Map position: 2
C;Superfamily: Saccharomyces cerevisiae probable membrane protein YMR100w

Query Match 27.4%; Score 57.5; DB 2; Length 574;
Best Local Similarity 41.2%; Pred. No. 7.3;
Matches 14; Conservative 6; Mismatches 11; Indels 3; Gaps 1;

Query Match 25.7%; Score 54; DB 2; Length 245;
Best Local Similarity 40.0%; Pred. No. 8.6;
Matches 12; Conservative 3; Mismatches 9; Indels 6; Gaps 1;

QY 8 RGTCDLWINEMSLHKIVQDVYGPDPHP 37
||| : : : : :
Db 73 RGRE-----ELKLAQWVFQDPYGLPHRH 96
||| : : : : :

RESULT 8
B1978
Hypothetical protein jhp0054 - Helicobacter pylori (strain J99)
C:Species: Helicobacter pylori
A:Variety: strain J99
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
C:Accession: B71978
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
; Ives, C.; Gibson, R.; Werberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;
Nature 397, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path-
A:Reference number: A71800; MUID:99120557; PMID:9923682
A:Accession: B71978
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-522 <ARN>
A:Cross-references: UNIPROT:Q9ZN06; GB:AE001445; GB:AE001439; NID:g4154559; PIDN:AAD00563
A:Experimental source: strain J99
C:Genetics:
A:Gene: jhp0054

Query Match 25.2%; Score 53; DB 2; Length 522;
Best Local Similarity 54.2%; Pred. No. 27;
Matches 13; Conservative 3; Mismatches 6; Indels 2; Gaps 1;

QY 2 AMLLDRRG--TECDLWINEMSLH 23
||| : : : : :
Db 338 AMLFDRNGVEFCDLKVCKELLN 361
||| : : : : :

RESULT 9
AF3338
Hypothetical protein BMEI0692 [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C:Accession: AF3338
R:Delvecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letese
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A:Reference number: AJ3252; PMID:11756688
A:Accession: AF3338
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-131 <KUR>
A:Cross-references: UNIPROT:Q8YHV4; GB:AE008917; PIDN:AAL51873.1; PID:g17982624; GSPDB:G
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BMEI0692
A:Map position: 1

Query Match 24.8%; Score 52; DB 2; Length 131;
Best Local Similarity 34.8%; Pred. No. 8;
Matches 8; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 13 DLWINEMSLHKIVQDVYGTGTP 35
: : : : :
Db 61 ELWAKVLGLVLAAPHLYGAPHP 83
: : : : :

RESULT 10
T08589
Hypothetical protein L23H3.40 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C:Accession: T08589
R:Bevan, M.; Pohl, T.; Weizenegger, T.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Schue
submitted to the Protein Sequence Database, May 1999
A:Reference number: Z16098
A:Accession: T08589
A:Molecule type: DNA
A:Residues: 1-297 <BEV>
A:Cross-references: UNIPROT:Q9SZ20; EMBL:AL050398; GSPDB:GN00062; ATSP:L23H3.40
A:Experimental source: cultivar Columbia; BAC clone L23H3
C:Genetics:
A:Gene: ATSP:L23H3.40
A:Map position: 4
A:Introns: 9/3; 44/3; 82/3; 142/3; 169/2; 196/2; 225/3; 252/3

Query Match 24.5%; Score 51.5; DB 2; Length 297;
Best Local Similarity 54.2%; Pred. No. 24;
Matches 13; Conservative 2; Mismatches 6; Indels 3; Gaps 1;

QY 13 DLWINEMSLHKI---VQDVYGTGTP 33
||| : : : : :
Db 210 DLMKETSLLNSIGSQFQDVIGTP 233
||| : : : : :

RESULT 11
A81272
Probable dCTP deaminase (EC 3.5.4.13) Cj1292 [imported] - Campylobacter jejuni (strain
C:Species: Campylobacter jejuni
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: A81272
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilli;
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barre
Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hy
A:Reference number: A81250; MUID:20150912; PMID:10688204
A:Accession: A81272
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-186 <PAR>
A:Cross-references: UNIPROT:Q9PN07; GB:AL139078; GB:AL111168; NID:g6968723; PIDN:CAB737
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
A:Gene: dcd; Cj1292
C:Superfamily: dCTP deaminase
C:Keywords: hydrolase

Query Match 24.3%; Score 51; DB 2; Length 186;
Best Local Similarity 42.1%; Pred. No. 16;
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 9 GTCCDLWINEMSLHKIVQ 27
||| : : : : :
Db 2 GLKADNWIRKMALEHKWIE 20
||| : : : : :

RESULT 12
A98302
Oligopeptide transport ATP-binding protein amfF AGR_L_2744 [imported] - Agrobacterium t
C:Species: Agrobacterium tumefaciens
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C:Accession: A98302
R:Goodner, B.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tu
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: A98302
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-259 <KUR>
A:Cross-references: UNIPROT:Q8UAB8; GB:AE007870; PIDN:AAK89939.1; PID:g15159894; GSPDB:
C:Genetics:

A:Gene: AGR_L 2744
A:Map position: linear chromosome

Query Match 24.3%; Score 51; DB 2; Length 259;
Best Local Similarity 33.3%; Pred. No. 24;
Matches 12; Conservative 5; Mismatches 19; Indels 0; Gaps 0;

QY 3 MLLDRRGTECDLWINEMSLHLKIVQDVYGTTPHPHS 38

DB 72 ILIDGRSVRDIPLIERCRMQLQWVQDPYGLHPQRS 107

RESULT 13

AG2981
Hypothetical protein Atu3456 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)

C:Species: Agrobacterium tumefaciens

C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004

C:Accession: AG2981

R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.B.; Chen, Y.; Woo, I.

; Karp, P.; Romero, P.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell

Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,

star, E.W.

A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; MUID:21608550; PMID:11743193

A:Accession: AG2981

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-259 <MUR>

A:Cross-references: UNIPROT:Q8UAB8; GB:AE008689; PIDN:AAL44269.1; PID:gl7741855; GSPDB:C

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: Atu3456

A:Map position: linear chromosome

Query Match 24.3%; Score 51; DB 2; Length 259;

Best Local Similarity 33.3%; Pred. No. 24;

Matches 12; Conservative 5; Mismatches 19; Indels 0; Gaps 0;

QY 3 MLLDRRGTECDLWINEMSLHLKIVQDVYGTTPHPHS 38

DB 72 ILIDGRSVRDIPLIERCRMQLQWVQDPYGLHPQRS 107

RESULT 14

G85930

Interacts with secY [imported] - Escherichia coli (strain O157:H7, substrain EDL933)

C:Species: Escherichia coli

C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004

C:Accession: G85930

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew

illier, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,

Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: G85930

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-181 <STO>

A:Cross-references: UNIPROT:Q8X6T2; GB:AE005174; NID:gl2517269; PIDN:AAG57907.1; GSPDB:C

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

A:Gene: syd

C:Superfamily: Escherichia coli syd protein

Query Match 24.0%; Score 50.5; DB 2; Length 181;

Best Local Similarity 34.2%; Pred. No. 19;

Matches 13; Conservative 4; Mismatches 14; Indels 7; Gaps 2;

QY 1 AAMLLDRRGTECDLWINEMSLHLK---IVQDVYGTTPHP 35

DB 6 AQALKDFTARYCDWHEE----HKSWPLSEELYGVPS 39

RESULT 15

E91085

Syd protein [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509952)

C:Species: Escherichia coli

C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004

C:Accession: E91085

R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.

gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen

A:Reference number: A99629; MUID:21156231; PMID:11258796

A:Accession: E91085

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-181 <HAY>

A:Cross-references: UNIPROT:Q8X6T2; GB:BA000007; PIDN:BA037076.1; PID:gl13363125; GSPDB:C

A:Experimental source: strain O157:H7, substrain RIMD 0509952

C:Genetics:

A:Gene: Ece3653

C:Superfamily: Escherichia coli syd protein

Query Match 24.0%; Score 50.5; DB 2; Length 181;

Best Local Similarity 34.2%; Pred. No. 19;

Matches 13; Conservative 4; Mismatches 14; Indels 7; Gaps 2;

QY 1 AAMLLDRRGTECDLWINEMSLHLK---IVQDVYGTTPHP 35

DB 6 AQALKDFTARYCDWHEE----HKSWPLSEELYGVPS 39

Search completed: November 10, 2004, 14:52:16

Job time : 10.3208 secs

Result No.	Query Match	Score	Length	DB	ID	Description	
1	67	31.9	634	2	Q6C6T9	Q6C6T9 yarrowia li	
2	61.5	29.3	167	2	Q93Q79	Q93Q79 rhodococcus	
3	58	27.6	537	2	Q6SGV8	Q6SGV8 uncultured	
4	58	27.6	537	2	AAS07882	AAS07882 uncultured	
5	58	27.6	1011	2	Q82N43	Q82N43 streptomyces	
6	57.5	27.4	308	2	Q9KCA5	Q9KCA5 bacillus ha	
7	57.5	27.4	574	1	YB2A_SCHPO	P87311 schizosacch	
8	57	27.1	184	2	O06269	O06269 mycobacteri	
9	57	27.1	184	2	Q7TWM35	Q7TWM35 mycobacteri	
10	57	27.1	302	2	Q81527	Q81527 plasmodium	
11	57	27.1	302	2	Q81723	Q81723 plasmodium	
12	57	27.1	430	2	Q961B9	Q961B9 drosophila	
13	57	27.1	1116	1	YK54_AQUAE	Q67838 aquifex ae	
14	57	27.1	1568	2	Q81332	Q81332 plasmodium	
15	56.5	26.9	503	2	Q81020	Q81020 mus musculu	
16	55.5	26.4	431	1	ARGD_BIFLO	Q93482 bifidobacte	
17	55	26.2	312	2	Q934B2	Q934B2 caenorhabdi	
18	55	26.2	1251	2	Q9K4E9	Q9K4E9 streptomyce	
19	54.5	26.0	101	2	Q9CYU9	Q9CYU9 mus musculu	
20	54.5	26.0	756	2	Q7M762	Q7M762 mus musculu	
21	54.5	26.0	1317	2	Q9GQC3	Q9GQC3 brugia mala	
22	54.5	26.0	1608	2	Q9FW76	Q9FW76 cryza sativ	
23	54	25.7	245	2	Q7CVF0	Q7CVF0 agrobacteri	
24	54	25.7	245	2	Q8U736	Q8U736 agrobacteri	
25	54	25.7	398	1	DKR_BUCAP	Q8K9S7 buchnera ap	
26	54	25.7	401	2	Q9FL51	Q9FL51 streptomyces	
27	54	25.7	484	2	Q93YK3	Q93YK3 streptococc	
28	54	25.7	486	2	Q8NZY7	Q8NZY7 streptococc	
29	54	25.7	517	1	CV04_HUMAN	Q8WU47 homo sapien	
30	54	25.7	517	2	CAG30298	CAG30298 homo sapi	
31	54	25.7	541	2	Q9N7D6	Q9N7D6 homo sapien	

01-DEC-2001 (TrEMBLrel. 19, Created)
 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 Benzocate dioxygenase small subunit.
 Name=bop1;
 Rhodococcus sp. 19070.
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Corynebacterineae; Nocardiaceae; Rhodococcus.
 NCBI_TaxID=161384;
 [1] SEQUENCE FROM N.A.
 RP STRAIN=19070;
 RC MEDLINE=21268842; PubMed=11375157;
 RA Haddad S., Eby D.M., Neidle E.L.;
 RT "Cloning and expression of the benzoate dioxygenase genes from
 Rhodococcus sp. strain 19070";
 RL Appl. Environ. Microbiol. 67:2507-2514 (2001).
 DR EMBL; AF273141; RAK58904.1; -
 DR GO; GO:0003824; Catalytic activity; IEA.
 DR GO; GO:0016702; Peroxidoreductase activity, acting on single d. . .; IEA.
 DR GO; GO:0006725; P-atomate compound metabolism; IEA.
 DR GO; GO:0006118; P-electron transport; IEA.
 DR InterPro; IPR000391; Ring hydroxyl B.
 DR Pfam; PF00866; Ring_hydroxyl_B; 1.
 KW Dioxygenase
 SQ SEQUENCE 167 AA; 19861 MW; CAD4BCD83D692D91 CRC64;
 Query Match 29.3%; Score 61.5; DB 2; Length 167;
 Best Local Similarity 38.6%; Pred. No. 2.4;
 Matches 17; Conservative 5; Mismatches 9; Indels 13; Gaps 2;
 QY 4 LLDRTGTECDLWNEMLLHK-----IVQDVVGTGPH 35
 DB 102 ILRRGTEDLRHNFSLRYINTDTYFGTNFVLDSG-PHP 144

RESULT 3
 Q6SGV8 PRELIMINARY; PRT; 537 AA.
 ID Q6SGV8
 AC Q6SGV8; (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=EBAC080-L32B05.10;
 OS uncultured bacterium 463.
 OC Bacteria; environmental samples.
 NCBI_TaxID=257394;
 [1] SEQUENCE FROM N.A.
 RP STRAIN=MA-4680;
 RC MEDLINE=21477403; PubMed=11572948;
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
 RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
 RT "Genome sequence of an industrial microorganism Streptomyces
 avermitilis: deducing the ability of producing secondary
 metabolites";
 RT Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
 [2] SEQUENCE FROM N.A.
 RP STRAIN=MA-4680;
 RC MEDLINE=22608306; PubMed=12692562;
 RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
 RA Sakaki Y., Hattori M., Omura S.;
 RT "Complete genome sequence and comparative analysis of the industrial
 microorganism Streptomyces avermitilis";
 RL Nat. Biotechnol. 21:528-531(2003).
 DR EMBL; AP005027; BAC69170.1; -
 DR InterPro; IPR001440; TPR.
 DR Pfam; PF00515; TPR; 4.
 DR SMART; SM00028; TPR; 4.
 DR PROSITE; PS02993; TPR_REGION: 1.
 KW Complete proteome; Repeat; TPR repeat.

01-DEC-2001 (TrEMBLrel. 19, Created)
 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 Benzocate dioxygenase small subunit.
 Name=bop1;
 Rhodococcus sp. 19070.
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Corynebacterineae; Nocardiaceae; Rhodococcus.
 NCBI_TaxID=161384;
 [1] SEQUENCE FROM N.A.
 RP STRAIN=19070;
 RC MEDLINE=21268842; PubMed=11375157;
 RA Haddad S., Eby D.M., Neidle E.L.;
 RT "Cloning and expression of the benzoate dioxygenase genes from
 Rhodococcus sp. strain 19070";
 RL Appl. Environ. Microbiol. 67:2507-2514 (2001).
 DR EMBL; AF273141; RAK58904.1; -
 DR GO; GO:0003824; Catalytic activity; IEA.
 DR GO; GO:0016702; Peroxidoreductase activity, acting on single d. . .; IEA.
 DR GO; GO:0006725; P-atomate compound metabolism; IEA.
 DR GO; GO:0006118; P-electron transport; IEA.
 DR InterPro; IPR000391; Ring hydroxyl B.
 DR Pfam; PF00866; Ring_hydroxyl_B; 1.
 KW Dioxygenase
 SQ SEQUENCE 167 AA; 19861 MW; CAD4BCD83D692D91 CRC64;
 Query Match 29.3%; Score 61.5; DB 2; Length 167;
 Best Local Similarity 38.6%; Pred. No. 2.4;
 Matches 17; Conservative 5; Mismatches 9; Indels 13; Gaps 2;
 QY 4 LLDRTGTECDLWNEMLLHK-----IVQDVVGTGPH 35
 DB 102 ILRRGTEDLRHNFSLRYINTDTYFGTNFVLDSG-PHP 144

RESULT 3
 Q6SGV8 PRELIMINARY; PRT; 537 AA.
 ID Q6SGV8
 AC Q6SGV8; (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=EBAC080-L32B05.10;
 OS uncultured bacterium 463.
 OC Bacteria; environmental samples.
 NCBI_TaxID=257394;
 [1] SEQUENCE FROM N.A.
 RP STRAIN=MA-4680;
 RC MEDLINE=21477403; PubMed=11572948;
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
 RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
 RT "Genome sequence of an industrial microorganism Streptomyces
 avermitilis: deducing the ability of producing secondary
 metabolites";
 RT Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
 [2] SEQUENCE FROM N.A.
 RP STRAIN=MA-4680;
 RC MEDLINE=22608306; PubMed=12692562;
 RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
 RA Sakaki Y., Hattori M., Omura S.;
 RT "Complete genome sequence and comparative analysis of the industrial
 microorganism Streptomyces avermitilis";
 RL Nat. Biotechnol. 21:528-531(2003).
 DR EMBL; AP005027; BAC69170.1; -
 DR InterPro; IPR001440; TPR.
 DR Pfam; PF00515; TPR; 4.
 DR SMART; SM00028; TPR; 4.
 DR PROSITE; PS02993; TPR_REGION: 1.
 KW Complete proteome; Repeat; TPR repeat.

```

SQ SEQUENCE 1011 AA; 111014 MW; 2ADBEA2562166AA9 CRC64;

Query Match
Best Local Similarity 27.6%; Score 58; DB 2; Length 1011;
Matches 13; Conservative 8; Mismatches 12; Indels 2; Gaps 1;

QY 1 AAMLLDRGTECDLWINEMSLHKKIVQDVYGTGTHP 35
DB 916 AHLLLDQGNSEAM--TEVNAHRIYTKYGRHP 948

RESULT 6
Q9KCA5 PRELIMINARY; PRT; 308 AA.
AC Q9KCA5
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE BH1668 protein.
GN Name=BH1668;
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
DR EMBL; AP001512; BAB05387.1; -.
DR PIR; D83858; D83858.
DR InterPro; IPR011059; Metallo hydrolase.
SQ SEQUENCE 308 AA; 35202 MW; 39DB871087B7DE77 CRC64;

Query Match
Best Local Similarity 27.4%; Score 57.5; DB 2; Length 308;
Matches 12; Conservative 5; Mismatches 4; Indels 5; Gaps 2;

QY 13 DLWINEMSLHKKIVQDVYGTGTHPHPS 38
DB 189 DLWVS-LCEAHR1----YGTTHPPNS 209

RESULT 7
YB2A_SCHPO STANDARD; PRT; 574 AA.
AC P87311,
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Hypothetical protein C31F10.10c in chromosome II.
GN ORFNames=SPBC31F10.10c;
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
GN NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360; DOI=10.1038/nature724;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Murgall K., Murphy L., Niblett D., Odell C.,

```

```

RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skellon J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vansreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehman H., Reinhardt R., Fohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Sapakowski G.V., Ussery D., Barrett B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -|- SUBCELLULAR LOCATION: MYND-type zinc finger.
CC -|- SIMILARITY: Contains 1 MYND-type zinc finger.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z97204; CAB10087.1; -.
CC PIR; T40213; T40213.
CC GeneDB Spombe; SPBC31F10.10c; -.
CC InterPro; IPR008938; ARM.
CC InterPro; IPR002893; Znf MYND.
CC Pfam; PF01753; Zf-MYND; 1.
CC PROSITE; PS01360; ZF-MYND_1; 1.
CC PROSITE; PS00865; ZF-MYND_2; 1.
CC DNA-binding; Hypothetical protein; Nuclear protein; Zinc-finger.
FW ZN FING 482 523 MYND-type.
SQ SEQUENCE 574 AA; 64142 MW; E668B23164D3F75 CRC64;

Query Match
Best Local Similarity 27.4%; Score 57.5; DB 1; Length 574;
Matches 14; Conservative 6; Mismatches 11; Indels 3; Gaps 1;

QY 3 MLLDRGTECDLWINEMSLHKKIVQDVYGTGTHP 36
DB 21 VLYDREALDCD---SEMSLWNSLSHLVLTSTSP 51

RESULT 8
O06269 PRELIMINARY; PRT; 184 AA.
AC O06269; Q7D573;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein MT3716.
GN OrderedLocustNames=MT3716, RV3614c;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
GN NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
RA Harris D.E., Gordon S.V., Eiglmeier K., Gas S., Barry C.E. II,
RA Tekala E., Badcock K., Basham D., Brown D., Chillingworth T.,
RA Connor R., Davies K.M., Devlin K., Feltwell T., Gentles S., Hamlin N.,
RA Holroyd S., Hornsby T., Jagsels K., Krogh A., McLean J., Moule S.,
RA Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,

```

```

RA Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RA "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
RA Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
RA Salzberg S.L., Delcher A., Uterback T.R., Weidman J.F., Khouri H.M.,
RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA Fraser C.M.;
RA "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RT J. Bacteriol. 184:5479-5490(2002).
DR EMBL; BX842583; CAB08952.1; -
DR EMBL; AB007171; AKA48075.1; -
DR FIR; G70556; G70956.
DR TIGR; WT3716; -
DR TubercuList; RV3614c; -
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 184 AA; 19834 MW; 6326911F65DC5C7A CRC64;

Query Match 27.1%; Score 57; DB 2; Length 184;
Best Local Similarity 35.3%; Pred. No. 12;
Matches 12; Conservative 7; Mismatches 15; Indels 0; Gaps 0;

QY 2 AMLDRRGTECDLWINEXSLHKIVQDVYGTGPH 35
Db 122 AFILDRMSQQVDADEHRAVALLAKTVGTWGLPSP 155

RESULT 9
Q7TW35 PRELIMINARY; PRT; 184 AA.
AC Q7TW35;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein Mb3644c.
GN OrderedLocusNames=Mb3644c;
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1765;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972;
RA Garnier T., Eiglmier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrall B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RA "The complete genome sequence of Mycobacterium bovis.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
DR EMBL; BX248346; CAD95830.1; -
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 184 AA; 19834 MW; 6326911F65DC5C7A CRC64;

Query Match 27.1%; Score 57; DB 2; Length 184;
Best Local Similarity 35.3%; Pred. No. 12;
Matches 12; Conservative 7; Mismatches 15; Indels 0; Gaps 0;

QY 2 AMLDRRGTECDLWINEXSLHKIVQDVYGTGPH 35
Db 122 AFILDRMSQQVDADEHRAVALLAKTVGTWGLPSP 155

RESULT 10
Q81527 PRELIMINARY; PRT; 302 AA.
AC Q81527;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hydroxyethylthiazole kinase, putative.
GN C8FNames=PFL1920c;
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=33329;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22255705; PubMed=12368864;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlon J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.S., Nene V., Shallow S.J., Suh B., Peterson J., Augiuoli S.,
RA Martea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.;
RA "Genome sequence of the human malaria parasite Plasmodium
RT falciparum.";
RT Nature 419:498-511(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Hyman R.W., Fung E., Conway A., Kurdi O., Mao J., Miranda M.,
RA Nakao B., Rowley D., Tanaki T., Wang F., Davis R.W.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB014850; AAN36470.1; -
DR HSSP; P39593; IEKQ.
DR GO; GO:0004417; F:hydroxyethylthiazole kinase activity; IEA.
DR GO; GO:0009228; P:thiamin biosynthesis; IEA.
DR InterPro; IPR000417; Hyethyz_kinase.
DR Pfam; PF02110; HK; 1.
DR PRINTS; PR01099; HYETHYZKNASE.
KW Kinase.
SQ SEQUENCE 302 AA; 33882 MW; C743CBD06B6245EE CRC64;

Query Match 27.1%; Score 57; DB 2; Length 302;
Best Local Similarity 55.6%; Pred. No. 20;
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 20 SLHKIVQDVYGTGPHPPH 37
Db 272 SLSHKIIDIYYSHNPH 289

RESULT 11
Q81723 PRELIMINARY; PRT; 302 AA.
AC Q81723;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE 4-methyl-5-beta-hydroxyethylthiazole kinase.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5633;
RN [1]
RP SEQUENCE FROM N.A.
RA Tami G., Pelle R., Mulaa P.J.;
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY166865; AAN86938.1; -
DR HSSP; P39593; IEKQ.
DR GO; GO:0004417; F:hydroxyethylthiazole kinase activity; IEA.
DR GO; GO:0009228; P:thiamin biosynthesis; IEA.
DR InterPro; IPR000417; Hyethyz_kinase.
DR Pfam; PF02110; HK; 1.
DR PRINTS; PR01099; HYETHYZKNASE.
KW Kinase.
SQ SEQUENCE 302 AA; 33868 MW; 638E08852FC867F3 CRC64;

```

```

Query Match          27.1%; Score 57; DB 2; Length 302;
Best Local Similarity 55.6%; Pred. No. 20;
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 20 SLHKIVQDVGTPPH 37
DB 272 SLHKIIDVYYSHNPH 289

RESULT 12
ID Q961B9 PRELIMINARY; PRT; 430 AA.
AC Q961B9; Q9VUZ7;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
DE LD24073p (CG5027-PA).
GN ORFNames=CG5027;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nuncio J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E., Heit G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoef C., Baldwin D.,
RA Baile R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.A., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jaiswal M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kamali M.B., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun B.,
RA Svirska R., Tector R., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Massarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodger, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhou S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
```

```

Science 287:2185-2195 (2000).
[3]
SEQUENCE FROM N.A.
MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirska R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstock G., Scher S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: release 3 of the Drosophila
RT melanogaster euchromatic genome sequence.";
```

Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).

[4]

SEQUENCE FROM N.A.

MEDLINE=22426070; PubMed=12537573;

RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirska R.,

R Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,

RA Ashburner M., Celniker S.E.;

RT "The transposable elements of the Drosophila melanogaster euchromatin:

RT a genomics perspective.";

Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).

[5]

SEQUENCE FROM N.A.

MEDLINE=22426069; PubMed=12537572;

RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,

RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,

RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,

RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,

RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,

RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,

RA Lewis S.E.;

RT "Annotation of the Drosophila melanogaster euchromatic genome: a

RT systematic review.";

Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).

[6]

SEQUENCE FROM N.A.

FLYBASE;

RP Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.

[7]

SEQUENCE FROM N.A.

FLYBASE;

RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: Calsequestrin is a high-capacity, moderate affinity,

CC calcium-binding protein and thus acts as an internal calcium store

CC in muscle. The release of calcium bound to calsequestrin through a

CC calcium release channel triggers muscle contraction (By

CC similarity).

CC -!- SIMILARITY: Belongs to the calsequestrin family.

DR EMBL; AY051696; RAK93120.1; -.

DR EMBL; AR003528; RAF49525.3; -.

DR HSSP; P07237; IMEX.

DR IntAct; Q961B9; -.

DR FlyBase; FBgn0036579; CG5027.

DR GO; GO:0005514; F:calcium ion storage activity; IEA.

DR GO; GO:0005489; F:electron transporter activity; IEA.

DR GO; GO:0006118; P:electron transport; IEA.

DR InterPro; IPR001393; Calsequestrin.

DR InterPro; IPR006662; Thioered.

DR InterPro; IPR006663; Thioeredox_dom2.

DR Pfam; PF00085; Thioeredoxin; 1.

DR PRINTS; PR00312; CALSEQUESTRN.

DR PRINTS; PR00421; THIOREDOXIN.

DR PROSITE; PS00194; THIOREDOXIN; 1.

KW Calcium; Calcium-binding; Redox-active center.

SK SEQUENCE 430 AA; 49507 MW; 782F66BF5607DF7 CRC64;

Query Match 27.1%; Score 57; DB 2; Length 430;

Best Local Similarity 50.0%; Pred. No. 31;

Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 15 WINEMSLHKIVQDVGTPH 34

|||||:|:|:|

Db 301 WIGPSIAHSIILDQLTPH 320

RESULT 13

YK54_AQUAE STANDARD; PRT; 1116 AA.

AC 067838;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DE Hypothetical protein AQ_2054.

DE Hypothetical protein AQ_2054.

GN OrderedLocusNames=AQ_2054;

OS Aquifex aeolicus.

OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.

OX NCBI_TaxID=63363;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=VF5;

RX MEDLINE=98196666; PubMed=9537320; DOI=10.1038/32831;

RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,

RA Graham D.E., Overbeek R., Snead M.A., Keller M., Aujaay M., Huber R.,

RA Feldman R.A., Short J.M., Olsen G.J., Swanson R.V.;

RT "The complete genome of the hyperthermophilic bacterium Aquifex

RT aeolicus.";

RL Nature 392:353-358(1998).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AF000770; AAC07805.1; -

DR PIR; B70476; B70476; DUF490.

DR InterPro; IPR007452; DUF490.

DR Pfam; PF04357; DUF490; 1.

KW Complete proteome; Hypothetical protein; Transmembrane.

FT TRANSMEM 3 20 Potential.

SQ SEQUENCE 1116 AA; 127794 MW; 423B5FCC6C02C89C CRC64;

Query Match 27.1%; Score 57; DB 1; Length 1116;

Best Local Similarity 41.9%; Pred. No. 93;

Matches 13; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

QY 6 DRRGTCECLINEMSLHKKIVQDVYGTPTTP 36

Db 950 DRELTEFDYFLNTSDGVKIFLLVHGTPTNP 980

RESULT 14

Q81332 PRELIMINARY; PRT; 1568 AA.

AC Q81332;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Hypothetical protein PF10585C.

GN Names=PF10585C;

OS Plasmodium falciparum (isolate 3D7).

OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.

OX NCBI_TaxID=36329;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=2255708; PubMed=12368867;

RA Hall N., Pain A., Berriman M., Churcher C., Harris B., Harris D.,

RA Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,

RA Buckee C.O., Burrows C., Cherevach I., Chillingworth C.,

RA Chillingworth T., Christodoulou Z., Clark L., Clark R., Corton C.,

RA Cronin A., Davies R., Davis P., Dear P., Dearden F., Doggett J.,

RA Feltwell T., Goble A., Goodhead I., Gwilliam R., Hamlin N., Hance Z.,

RA Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P.,

RA Humphray S., Jagels K., James K.D., Johnson D., Kerhornou A.,

RA Knights A., Konfortov B., Kyes S., Larke N., Lawson D., Lennard N.,

RA Line A., Maddison M., McLean J., Mooney P., Moule S., Murphy L.,

RA Oliver K., Ormond D., Price C., Quail M.A., Rabinowitsch E.,

RA Rajandream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M.,

RA Seeger K., Sharp S., Smith R., Squares R., Squares S., Stevens K.,

RA Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,

RA Sulston J.E., Craig A., Newbold C., Barrell B.G.;

RT "Sequence of plasmodium falciparum chromosomes 1, 3-9 and 13.";

RL Nature 419:527-531(2002).

DR EMBL; AL929356; CAD51803.1; -

KW Hypothetical protein.

SQ SEQUENCE 1568 AA; 188685 MW; FLD966E66CB78DE7 CRC64;

Query Match 27.1%; Score 57; DB 2; Length 1568;

Best Local Similarity 48.0%; Pred. No. 1.4e+02;

Matches 12; Conservative 5; Mismatches 4; Indels 4; Gaps 1;

QY 10 TSCD----LWINEMSLHKKIVQDVY 30

Db 958 TLCDILINLYINKMDVLKILNDIY 982

RESULT 15

Q81020 PRELIMINARY; PRT; 503 AA.

AC Q81020;

DT 01-JUN-2003 (TrEMBLrel. 24, Created)

DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE PRAMEL4.

GN Name=Pramel4;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=22507960; PubMed=12620990;

RA Bortvin A., Egan K., Skaletsky H., Akutsu H., Berry D.L.,

RA Yanagimachi R., Page D.C., Jaenisch R.;

RT "Incomplete reactivation of Oct4-related genes in mouse embryos cloned

RT from somatic nuclei.";

RL Development 130:1673-1680(2003).

DR EMBL; AF490340; AAO84498.1; -

DR MGD; MGI:2156377; Pramel4.

SQ SEQUENCE 503 AA; 58137 MW; ELA8C0B005212B30 CRC64;

Query Match 26.9%; Score 56.5; DB 2; Length 503;

Best Local Similarity 39.4%; Pred. No. 44;

Matches 13; Conservative 4; Mismatches 15; Indels 1; Gaps 1;

QY 4 LLDRRCETCLATNEMSLHKKIVQDVYGTPTTP 35

Db 120 VLDRDAHQDFWDGWAGLLHEVCSDVFGKQF 152

Search completed: November 10, 2004, 14:50:02

Job time : 52.6189 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:44:14 , Search time 8.05283 Seconds
(without alignments)
181.178 Million cell updates/sec

Title: US-10-092-750-47

Perfect score: 124

Sequence: 1 PQYKPIADLYRGRESRPSAPR 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA.*
- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
 - 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
 - 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
 - 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
 - 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
 - 6: /cgn2_6/ptodata/1/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	40.3	293	4	US-09-540-236-2400
2	48	38.7	650	4	US-09-252-991A-19052
3	48	38.7	851	4	US-09-252-991A-19645
4	47	37.9	151	4	US-09-621-976-4254
5	47	37.9	297	4	US-09-252-991A-18764
6	47	37.9	502	4	US-10-140-002-548
7	46	37.1	186	4	US-09-252-991A-22673
8	46	37.1	285	4	US-09-252-991A-32954
9	46	37.1	372	4	US-09-252-991A-27599
10	46	37.1	458	4	US-09-252-991A-28585
11	46	37.1	724	4	US-09-562-737-22
12	46	37.1	724	4	US-09-562-737-22
13	45.5	36.7	204	4	US-09-513-999C-7997
14	45.5	36.7	204	4	US-09-513-999C-7998
15	45.5	36.7	219	3	US-08-924-747-4
16	45.5	36.7	219	3	US-09-247-373B-4
17	45.5	36.7	219	3	US-08-296-715-4
18	45.5	36.7	326	4	US-09-252-991A-27158
19	45.5	36.7	1220	2	US-08-843-530B-36
20	45.5	36.7	1220	4	US-09-636-728-32
21	45	36.3	182	4	US-09-252-991A-27971
22	45	36.3	208	4	US-09-538-092-810
23	45	36.3	724	4	US-09-562-737-27
24	45	36.3	724	4	US-08-582-737-30
25	44.5	35.9	904	4	US-09-252-991A-19257
26	44	35.5	79	4	US-09-489-039A-8889
27	44	35.5	190	4	US-09-252-991A-25320

ALIGNMENTS

RESULT 1

US-09-540-236-2400
; Sequence 2400, Application US/09540236
; Patent No. 6673910

GENERAL INFORMATION:

- APPLICANT: Gary L. Breton et al.
- TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CAT.
- TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
- FILE REFERENCE: 2709.2005-001
- CURRENT APPLICATION NUMBER: US/09/540,236
- CURRENT FILING DATE: 2000-04-04
- NUMBER OF SEQ ID NOS: 3840
- SEQ ID NO 2400
- LENGTH: 293
- TYPE: PRT
- ORGANISM: M.catarrhalis

Query Match 40.3%; Score 50; DB 4; Length 293;

Best Local Similarity 44.4%; Pred. No. 9.8;

Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Query

2 WQYKPIADLYRGRESRPS 19

217 WSYMPVADRYLAQISNPA 234

RESULT 2

US-09-252-991A-19052
; Sequence 19052, Application US/09252991A
; Patent No. 6551795

GENERAL INFORMATION:

- APPLICANT: Marc J. Rubenfield et al.
- TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
- TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
- FILE REFERENCE: 107196.136
- CURRENT APPLICATION NUMBER: US/09/252,991A
- CURRENT FILING DATE: 1999-02-18
- PRIOR APPLICATION NUMBER: US 60/074,788
- PRIOR FILING DATE: 1998-02-18
- PRIOR APPLICATION NUMBER: US 60/094,190
- PRIOR FILING DATE: 1998-07-27
- NUMBER OF SEQ ID NOS: 33142
- SEQ ID NO 19052
- LENGTH: 650
- TYPE: PRT
- ORGANISM: Pseudomonas aeruginosa

Query Match 38.7%; Score 48; DB 4; Length 650;

Best Local Similarity 66.7%; Pred. No. 46;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 8 ADLYRGRESRPSAPR 22
DB 436 ADLGRGPRSRPARP 450

RESULT 3
US-09-252-991A-19645
; Sequence 19645, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19645
; LENGTH: 851
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19645

Query Match 38.7%; Score 48; DB 4; Length 851;
Best Local Similarity 40.9%; Pred. No. 61;
Matches 9; Conservative 4; Mismatches 7; Indels 2; Gaps 1;

QY 2 WQYKPIADLYRGSE--SRPSAP 21
DB 16 WPFRLAQDFNRNSWFLRPGAP 37

RESULT 4
US-09-621-976-4254
; Sequence 4254, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 4254
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 6
; OTHER INFORMATION: Xaa = Phe,Ile,Leu,Val
US-09-621-976-4254

Query Match 37.9%; Score 47; DB 4; Length 151;
Best Local Similarity 47.1%; Pred. No. 13;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 PWQYKPIADLYRGRESR 17
DB 84 PWESKNSTAVWRGRDSR 100

RESULT 5

US-09-252-991A-18764
; Sequence 18764, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18764
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18764

Query Match 37.9%; Score 47; DB 4; Length 297;
Best Local Similarity 47.1%; Pred. No. 28;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 PWQYKPIADLYRGRESR 17
DB 221 PWQGLPVAQRYSGSQCR 237

RESULT 6
US-10-140-002-548
; Sequence 548, Application US/10140002
; Patent No. 6725730
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330RIC59
; CURRENT APPLICATION NUMBER: US/10/140,002
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 548
; LENGTH: 502
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-002-548

Query Match 37.9%; Score 47; DB 4; Length 502;
Best Local Similarity 47.1%; Pred. No. 49;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 PWQYKPIADLYRGRESR 17
DB 297 PWESKNSTAVWRGRDSR 313


```

; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27599
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27599

Query Match 37.1%; Score 46; DB 4; Length 372;
Best Local Similarity 40.9%; Pred. No. 50;
Matches 9; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 1 PWQYKPTADLYRGRESRPSAPR 22
   |:::| | | | | | | | | | | |
Db 61 PCWRPNAGTGRARPTAPAPR 82

RESULT 10
US-09-252-991A-28585
; Sequence 28585, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28585
; LENGTH: 458
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28585

Query Match 37.1%; Score 46; DB 4; Length 458;
Best Local Similarity 52.9%; Pred. No. 62;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 6 PIADLYRGRESRPSAPR 22
   |:::| | | | | | | | | | | |
Db 61 PIRDPYRPIRPIRPTPR 77

RESULT 11
US-09-562-737-22
; Sequence 22, Application US/09562737
; Patent No. 6428967
; GENERAL INFORMATION:
; APPLICANT: Herz, Joachim
; TITLE OF INVENTION: LDL Receptor Signaling Pathways
; FILE REFERENCE: UTSW0708
; CURRENT APPLICATION NUMBER: US/09/562,737
; CURRENT FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 724
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Sequence
US-09-562-737-22

```

Query Match 37.1%; Score 46; DB 4; Length 724;
Best Local Similarity 42.9%; Pred. No. 1e+02;
Matches 9; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 1 PWQKPIADLYRGRESRPSAP 21
DB 291 PORYSPVAKURGEEDIPREP 311

RESULT 12
US-09-562-737-29
; Sequence 29, Application US/09562737
; Patent No. 6428967
; GENERAL INFORMATION:
; APPLICANT: Herz, Joachim
; APPLICANT: Gotthardt, Michael
; TITLE OF INVENTION: LDL Receptor Signaling Pathways
; FILE REFERENCE: UTSW0708
; CURRENT APPLICATION NUMBER: US/09/562,737
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 29
; LENGTH: 724
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Sequence
US-09-562-737-29

Query Match 37.1%; Score 46; DB 4; Length 724;
Best Local Similarity 40.9%; Pred. No. 1e+02;
Matches 9; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 PWQKPIADLYRGRESRPSAP 22
DB 291 PRRYSPVADLLGEEDIPTEPR 312

RESULT 13
US-09-513-999C-7997
; Sequence 7997, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A. J.Y.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 7997
; LENGTH: 204
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-513-999C-7997

Query Match 36.7%; Score 45.5; DB 4; Length 204;
Best Local Similarity 40.9%; Pred. No. 31;
Matches 9; Conservative 5; Mismatches 7; Indels 1; Gaps 1;

QY 2 WQYKPIADLYRG-RESRPSAPR 22
DB 28 WQYQLSALHRAPRTRPDKAR 49

RESULT 14
US-09-513-999C-7998
; Sequence 7998, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A. J.Y.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 7998
; LENGTH: 204
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-513-999C-7998

Query Match 36.7%; Score 45.5; DB 4; Length 204;
Best Local Similarity 40.9%; Pred. No. 31;
Matches 9; Conservative 5; Mismatches 7; Indels 1; Gaps 1;

QY 2 WQYKPIADLYRG-RESRPSAPR 22
DB 28 WQYQLSALHRAPRTRPDKAR 49

RESULT 15
US-08-924-747-4
; Sequence 4, Application US/08924747
; Patent No. 6083570
; GENERAL INFORMATION:
; APPLICANT: MCGONIGLE, BRIAN
; APPLICANT: O'KEEFE, DANIEL
; TITLE OF INVENTION: SOYBEAN GLUTATHIONE-S-TRANSFERASE
; TITLE OF INVENTION: ENZYMES
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E.I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WORD FOR WINDOWS 95
; SOFTWARE: MICROSOFT WORD VERSION 7.0A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/924,747
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: FLOYD, LINDA AXAMETHY
; REGISTRATION NUMBER: 33,692
; REFERENCE/DOCKET NUMBER: CL-1108
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-892-8112
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 219 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:

```
;      TISSUE TYPE: SOYBEAN
;      IMMEDIATE SOURCE:
;      CLONE: SSM.PR0026.G11
US-08-924-747-4

Query Match          36.7%; Score 45.5; DB 3; Length 219;
Best Local Similarity 47.1%; Pred. No. 34;
Matches 8; Conservative 6; Mismatches 2; Indels 1; Gaps 1;
```

This Page Blank (usp10)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2004, 01:28:30 : Search time 25.4453 Seconds
(without alignments)
305.399 Million cell updates/sec

Title: US-10-092-750-47

Perfect score: 124

Sequence: 1 PWQVKPIADLYRGRESRPSAPR 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/ECT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/ECTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	124	100.0	22	14	US-10-092-750-47
2	55	44.4	61	17	US-10-425-115-333161
3	50	40.3	441	14	US-10-156-761-13264
4	50	40.3	445	15	US-10-417-700A-7
5	49	39.5	147	16	US-10-437-963-11116
6	49	39.5	552	14	US-10-369-493-15960
7	49	39.5	552	14	US-10-369-493-16323
8	49	39.5	555	14	US-10-369-493-15584
9	49	39.5	560	14	US-10-094-749-3278
10	49	39.5	901	14	US-10-120-801-71
11	49	39.5	1163	14	US-10-120-801-70
12	48	38.7	84	17	US-10-425-115-274406
13	48	38.7	212	15	US-10-320-797-3026

14	48	38.7	593	14	US-10-369-493-3763	Sequence 3763, App
15	47	37.9	125	16	US-10-437-963-187550	Sequence 187550, App
16	47	37.9	236	15	US-10-276-774-1997	Sequence 1997, App
17	47	37.9	311	17	US-10-425-115-274476	Sequence 274476, App
18	47	37.9	502	10	US-09-931-836-40	Sequence 40, Appl
19	47	37.9	502	10	US-09-746-783-148	Sequence 148, App
20	47	37.9	502	13	US-10-036-342-40	Sequence 40, Appl
21	47	37.9	502	13	US-10-036-041-40	Sequence 40, Appl
22	47	37.9	502	14	US-10-028-072-548	Sequence 548, App
23	47	37.9	502	14	US-10-035-855-40	Sequence 40, Appl
24	47	37.9	502	14	US-10-140-808-548	Sequence 548, App
25	47	37.9	502	14	US-10-121-049-548	Sequence 548, App
26	47	37.9	502	14	US-10-123-904-548	Sequence 548, App
27	47	37.9	502	14	US-10-140-470-548	Sequence 548, App
28	47	37.9	502	14	US-10-175-746-548	Sequence 548, App
29	47	37.9	502	14	US-10-176-918-548	Sequence 548, App
30	47	37.9	502	14	US-10-176-921-548	Sequence 548, App
31	47	37.9	502	14	US-10-036-214-40	Sequence 40, Appl
32	47	37.9	502	14	US-10-137-865-548	Sequence 548, App
33	47	37.9	502	14	US-10-140-474-548	Sequence 548, App
34	47	37.9	502	14	US-10-035-719-40	Sequence 40, Appl
35	47	37.9	502	14	US-10-142-431-548	Sequence 548, App
36	47	37.9	502	14	US-10-143-114-548	Sequence 548, App
37	47	37.9	502	14	US-10-140-002-548	Sequence 548, App
38	47	37.9	502	14	US-10-036-160-40	Sequence 40, Appl
39	47	37.9	502	14	US-10-142-419-548	Sequence 548, App
40	47	37.9	502	14	US-10-035-958-40	Sequence 40, Appl
41	47	37.9	502	14	US-10-036-150-40	Sequence 40, Appl
42	47	37.9	502	14	US-10-123-262-548	Sequence 548, App
43	47	37.9	502	14	US-10-142-423-548	Sequence 548, App
44	47	37.9	502	14	US-10-121-050-548	Sequence 548, App
45	47	37.9	502	14	US-10-141-755-548	Sequence 548, App

ALIGNMENTS

RESULT 1
US-10-092-750-47
; Sequence 47, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
; APPLICANT: Hamdon, Philip W.
; APPLICANT: Alpin, Julia
; APPLICANT: Wright, Martin C.
; TITLE OF INVENTION: Polypeptides Interactive with BCL-X1
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092,750
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/274,526
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 47
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-47

Query Match 100.0%; Score 124; DB 14; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.7e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PWQVKPIADLYRGRESRPSAPR 22

Db 1 PWQVKPIADLYRGRESRPSAPR 22

RESULT 2
US-10-425-115-333161
; Sequence 333161, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:

```
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 333161
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_66957C.1.pap
US-10-425-115-333161

Query Match      44.4%; Score 55; DB 17; Length 61;
Best Local Similarity 42.9%; Pred. No. 2.8;
Matches 9; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

Qy 2 WQYKPIADLYRGRESRPSAPR 22
Db 5 WQAHPPKLFYKGDGRPKDPR 25

RESULT 3
US-10-156-761-13264
; Sequence 13264, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 13264
; LENGTH: 441
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-13264

Query Match      40.3%; Score 50; DB 14; Length 441;
Best Local Similarity 64.3%; Pred. No. 99;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 8 ADLYRGRESRPSAP 21
Db 167 AELYGRDNDSPSAP 180

RESULT 4
US-10-417-700A-7
; Sequence 7, Application US/10417700A
; Publication No. US20040033581A1
; GENERAL INFORMATION:
; APPLICANT: ECOPIA BIOSCIENCES INC.
; APPLICANT: ZAZOPOULOS, Emmanuel
; APPLICANT: STAPPA, Alfredo
; APPLICANT: FARNET, Chris
; TITLE OF INVENTION: Specialized dual condensation/epimerization domain in non-ribosomal
```

```
; TITLE OF INVENTION: peptide synthetase systems
; FILE REFERENCE: 3002-14US
; CURRENT APPLICATION NUMBER: US/10/417,700A
; CURRENT FILING DATE: 2003-04-17
; NUMBER OF SEQ ID NOS: 139
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 7
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Actinoplanes sp.
US-10-417-700A-7

Query Match      40.3%; Score 50; DB 15; Length 445;
Best Local Similarity 58.8%; Pred. No. 1e+02;
Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 6 PIADLYRGRESRPSAPR 22
Db 410 PVADVPRGRVSRDPAAGR 426

RESULT 5
US-10-437-963-111116
; Sequence 11116, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 111116
; LENGTH: 147
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(147)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_15123C.1.pap
US-10-437-963-111116

Query Match      39.5%; Score 49; DB 16; Length 147;
Best Local Similarity 40.9%; Pred. No. 47;
Matches 9; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Qy 1 PWQKPIADLYRGRESRPSAPR 22
Db 22 PMLWQPTSSLHLVLRPPPPAPR 43

RESULT 6
US-10-369-493-15960
; Sequence 15960, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfang
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
```

```
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 15960
; LENGTH: 552
; TYPE: PRT
; ORGANISM: Xanthomonas campestris
US-10-369-493-15960

Query Match      39.5%; Score 49; DB 14; Length 552;
Best Local Similarity 62.5%; Pred. NO. 1.7e+02;
Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY      5 KPIADLYRGRESRPSA 20
Db      436 QPALDPYRGREISPSA 451

RESULT 7
US-10-369-493-16323
; Sequence 16323, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 16323
; LENGTH: 552
; TYPE: PRT
; ORGANISM: Xanthomonas campestris
US-10-369-493-16323

Query Match      39.5%; Score 49; DB 14; Length 552;
Best Local Similarity 62.5%; Pred. NO. 1.7e+02;
Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY      5 KPIADLYRGRESRPSA 20
Db      436 QPALDPYRGREISPSA 451

RESULT 8
US-10-369-493-15584
; Sequence 15584, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 15584

; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 15960
; LENGTH: 552
; TYPE: PRT
; ORGANISM: Xanthomonas campestris
US-10-369-493-15960

Query Match      39.5%; Score 49; DB 14; Length 552;
Best Local Similarity 62.5%; Pred. NO. 1.7e+02;
Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY      5 KPIADLYRGRESRPSA 20
Db      436 QPALDPYRGREISPSA 451

RESULT 9
US-10-094-749-3278
; Sequence 3278, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 3278
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-3278

Query Match      39.5%; Score 49; DB 14; Length 560;
Best Local Similarity 40.0%; Pred. NO. 1.7e+02;
Matches 8; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY      2 WOYKPIADLYRGRESRPSAP 21
Db      19 YERKPLSSVIRPLSRPEEP 38

RESULT 10
US-10-120-801-71
; Sequence 71, Application US/10120801
; Publication No. US20030203843A1
; GENERAL INFORMATION:
; APPLICANT: Pena, Carol
; APPLICANT: Guo, Xiaojia
; APPLICANT: Shimkets, Richard
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Spytek, Kimberly
```

```
; APPLICANT: Mehraban, Fuad
; APPLICANT: Topper, James N.
; APPLICANT: Malyankar, Uriel
; APPLICANT: Wasserman, Scott
; APPLICANT: Edinger, Shlomit
; APPLICANT: Smithson, Glennda
; APPLICANT: Gunther, Erik
; APPLICANT: Komuves, Laszlo
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-340
; CURRENT APPLICATION NUMBER: US/10/120,801
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 60/285748
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: 60/286068
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: 60/286292
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 60/288334
; PRIOR FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: 60/291241
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 60/322284
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: 60/285609
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 71
; LENGTH: 901
; TYPE: PRT
; ORGANISM: human
US-10-120-801-71

Query Match 39.5%; Score 49; DB 14; Length 901;
Best Local Similarity 42.9%; Pred. No. 2.8e+02;
Matches 9; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

Qy 1 PWQYKPIADLYRGSRPSAP 21
Db 873 PFKLSPTSPVYRGSSGSPSP 893

RESULT 11
US-10-120-801-70
; Sequence 70, Application US/10120801
; Publication No. US20030203843A1
; GENERAL INFORMATION:
; APPLICANT: Pena, Carol
; APPLICANT: Guo, Xiaojia
; APPLICANT: Shimkets, Richard
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Mehraban, Fuad
; APPLICANT: Topper, James N.
; APPLICANT: Malyankar, Uriel
; APPLICANT: Wasserman, Scott
; APPLICANT: Edinger, Shlomit
; APPLICANT: Smithson, Glennda
; APPLICANT: Gunther, Erik
; APPLICANT: Komuves, Laszlo
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-340
; CURRENT APPLICATION NUMBER: US/10/120,801
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 60/285748
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: 60/286068
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: 60/286292
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 60/288334
```

```
; PRIOR FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: 60/291241
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 60/322284
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: 60/285609
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 70
; LENGTH: 1163
; TYPE: PRT
; ORGANISM: human
US-10-120-801-70

Query Match 39.5%; Score 49; DB 14; Length 1163;
Best Local Similarity 42.9%; Pred. No. 3.6e+02;
Matches 9; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

Qy 1 PWQYKPIADLYRGSRPSAP 21
Db 1135 PFKLSPTSPVYRGSSGSPSP 1155

RESULT 12
US-10-425-115-274406
; Sequence 274406, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovacic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 274406
; LENGTH: 84
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_181849C.1.pap
US-10-425-115-274406

Query Match 38.7%; Score 48; DB 17; Length 84;
Best Local Similarity 50.0%; Pred. No. 37;
Matches 11; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

Qy 1 PWQYKPIADLYRGSRPSAPR 22
Db 35 PHQRAAPAKRTGCRSLPSSPR 56

RESULT 13
US-10-320-797-3026
; Sequence 3026, Application US/10320797
; Publication No. US20040014955A1
; GENERAL INFORMATION:
; APPLICANT: Broshkin, Alexey M.
; APPLICANT: Zamudio, Carlos
; TITLE OF INVENTION: IDENTIFICATION OF ESSENTIAL GENES OF CRYPTOCOCCUS NEOFORMANS AND
; TITLE OF INVENTION: METHODS OF USE
; FILE REFERENCE: 10182-021-999
; CURRENT APPLICATION NUMBER: US/10/320,797
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: 60/341,261
; PRIOR FILING DATE: 2001-12-17
; NUMBER OF SEQ ID NOS: 3361
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 3026
```



```

; LENGTH: 212
; TYPE: PRT
; ORGANISM: Cryptococcus neoformans
; US-10-320-797-3026

Query Match      38.7%; Score 48; DB 15; Length 212;
Best Local Similarity 45.0%; Pred. No. 92;
Matches 9; Conservative 5; Mismatches 4; Indels 2; Gaps 1;

QY 2 WQYKPIADLYRGRESRPSAP 21
Db 28 WEYRQLAVIHRA--SRPSRP 45

RESULT 14
US-10-369-493-3763
; Sequence 3763, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 3763
; LENGTH: 593
; TYPE: PRT
; ORGANISM: Neurospora crassa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(593)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-3763

Query Match      38.7%; Score 48; DB 14; Length 593;
Best Local Similarity 55.6%; Pred. No. 2.5e+02;
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 5 KPIADLYRGRESRPSAPR 22
Db 94 KPIFEPQGNESPPFAPR 111

RESULT 15
US-10-437-963-187550
; Sequence 187550, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 187550
; LENGTH: 125
; TYPE: PRT

```

```

; LENGTH: 212
; TYPE: PRT
; ORGANISM: Cryptococcus neoformans
; US-10-320-797-3026

Query Match      38.7%; Score 48; DB 15; Length 212;
Best Local Similarity 45.0%; Pred. No. 92;
Matches 9; Conservative 5; Mismatches 4; Indels 2; Gaps 1;

QY 2 WQYKPIADLYRGRESRPSAP 21
Db 28 WEYRQLAVIHRA--SRPSRP 45

RESULT 14
US-10-369-493-3763
; Sequence 3763, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 3763
; LENGTH: 593
; TYPE: PRT
; ORGANISM: Neurospora crassa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(593)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-3763

Query Match      38.7%; Score 48; DB 14; Length 593;
Best Local Similarity 55.6%; Pred. No. 2.5e+02;
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 5 KPIADLYRGRESRPSAPR 22
Db 94 KPIFEPQGNESPPFAPR 111

RESULT 15
US-10-437-963-187550
; Sequence 187550, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 187550
; LENGTH: 125
; TYPE: PRT

```

```

; LENGTH: 212
; TYPE: PRT
; ORGANISM: Cryptococcus neoformans
; US-10-320-797-3026

Query Match      38.7%; Score 48; DB 15; Length 212;
Best Local Similarity 45.0%; Pred. No. 92;
Matches 9; Conservative 5; Mismatches 4; Indels 2; Gaps 1;

QY 2 WQYKPIADLYRGRESRPSAP 21
Db 28 WEYRQLAVIHRA--SRPSRP 45

RESULT 14
US-10-369-493-3763
; Sequence 3763, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 3763
; LENGTH: 593
; TYPE: PRT
; ORGANISM: Neurospora crassa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(593)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-3763

Query Match      38.7%; Score 48; DB 14; Length 593;
Best Local Similarity 55.6%; Pred. No. 2.5e+02;
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 5 KPIADLYRGRESRPSAPR 22
Db 94 KPIFEPQGNESPPFAPR 111

RESULT 15
US-10-437-963-187550
; Sequence 187550, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 187550
; LENGTH: 125
; TYPE: PRT

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:40:53 ; Search time 5.39623 Seconds
(without alignments)
392.268 Million cell updates/sec

Title: US-10-092-750-47
Perfect score: 124
Sequence: 1 PWQVKPIADLYRGRESRPSAPR 22
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : FIR 79:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54	43.5	187	2 F72665	hypothetical prote
2	50.5	40.7	488	2 S74629	hypothetical prote
3	50	40.3	217	2 A84806	probable retrolem
4	50	40.3	282	2 A85076	hypothetical prote
5	48.5	39.1	204	2 S26380	ribosomal protein
6	48	38.7	432	2 T16141	hypothetical prote
7	47	37.9	1016	2 T31343	proline dehydrogen
8	46.5	37.5	323	2 T52297	squamosa promoter
9	46	37.1	213	2 E70633	hypothetical prote
10	46	37.1	290	2 A82650	transcription regu
11	46	37.1	290	2 G37432	probable transcrip
12	46	37.1	632	2 C71327	probable phosphogl
13	46	37.1	679	2 H84516	hypothetical prote
14	46	37.1	801	2 H83737	glucosidase BH0704
15	45.5	36.7	204	1 UC2369	ribosomal protein
16	45.5	36.7	306	2 H72417	conserved hypothet
17	45.5	36.7	1220	2 S48387	SN1 protein - yea
18	45	36.3	208	2 S59772	hypothetical prote
19	45	36.3	247	2 S16869	gene 147 protein -
20	45	36.3	295	2 S58285	alpha-soluble NSF
21	45	36.3	295	2 G02238	alpha-SNAP - human
22	45	36.3	295	2 S32367	alpha-SNAP protein
23	45	36.3	298	2 S32368	beta-SNAP protein
24	45	36.3	567	2 AC0143	choline dehydrogen
25	45	36.3	574	2 T29615	hypothetical prote
26	44.5	35.9	869	2 H83500	probable glucosyl
27	44	35.5	168	2 A52759	ECF family sigma f
28	44	35.5	177	2 C97540	ecf sigma factor (
29	44	35.5	312	2 A61183	hypothetical prote

30 35.5 467 2 E91112 hypothetical prote
31 35.5 467 2 G85957 unknown protein en
32 35.5 512 2 D87366 conserved hypothet
33 35.5 537 2 C90796 hypothetical prote
34 35.5 537 2 G85656 unknown in ISCs
35 35.5 536 1 S10901 choline dehydrogen
36 35.5 562 2 A85524 choline dehydrogen
37 35.5 562 2 E90673 choline dehydrogen
38 35.5 4450 2 JX0340 gramacidin S synth
39 35.5 4452 1 YGBSG2 gramacidin S synth
40 35.5 1509 1 A27224 myosin heavy chain
41 34.7 145 2 C72668 hypothetical prote
42 34.7 388 2 T27645 hypothetical prote
43 34.7 406 2 A90985 hypothetical prote
44 34.7 406 2 D85830 hypothetical prote
45 34.7 406 2 C64970 hypothetical prote

ALIGNMENTS

RESULT 1

F72665
hypothetical protein APE0749 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: F72665
R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takeawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.;
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aerop.
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: F72665
A;Status: Preliminary
A;Molecule type: DNA
A;Residues: 1-187 <XAW>
A;Cross-references: UNIPROT:Q9YE21; DBU:AP000060; NID:G5104189; PIDN:BAA79726.1; PID:
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE0749

Query Match 43.5%; Score 54; DB 2; Length 187;
Best Local Similarity 57.9%; Pred. No. 0.91;
Matches 11; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 4 YKPIADLYRGRESRPSAPR 22
|||:|||||||
Db 35 YKLVGLLYGLRELKPLPR 53

RESULT 2

S74629
hypothetical protein sll1686 - Synechocystis sp. (strain PCC 6803)
C;Species: Synechocystis sp.
A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C;Accession: S74629
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.
o, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasu
DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocyst:
s.
A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S74629
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-488 <KAN>
A;Cross-references: UNIPROT:P72766; EMBL:D90900; GB:AB001339; NID:G1651768; PIDN:BAA16
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 40.7%; Score 50.5; DB 2; Length 488;
Best Local Similarity 44.4%; Pred. No. 8.6;
Matches 8; Conservative 5; Mismatches 4; Indels 1; Gaps 1;

QY 1 PWQYKPIADLYRGRESRP 18
 |||||: : || :
 DB 130 PWQYRPVS-IQLGRRPKP 146

RESULT 3

A94806
 A;Title: probable retroelement pol polyprotein [imported] - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
 C;Accession: A84806
 R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, L.; Tallon, L.;
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402, 761-768, 1999
 A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A;Reference number: A84420; MUID:20083487; PMID:10617197
 A;Accession: A84806
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-217 <STO>
 A;Cross-references: UNIPROT:Q9ZVH6; GB:AE002093; NID:G3786020; PIDN:AA67366.1; GSPDB:GN
 C;Genetics:
 A;Gene: At2g38520
 A;Map position: 2

Query Match 40.3%; Score 50; DB 2; Length 217;
 Best Local Similarity 47.4%; Pred. No. 4.3; Mismatches 7; Indels 0; Gaps 0;
 Matches 9; Conservative 3;

QY 1 PWQYKPIADLYRGRESRPS 19
 |||||: ||||: :
 DB 178 PEDYKPIVDQMEGRTPPN 196

RESULT 4

A85076
 A;Title: Hypothetical protein At4g07710 [imported] - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
 C;Accession: A85076
 R;Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
 Nature 402, 769-777, 1999
 A;Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
 A;Reference number: A85001; MUID:20083488; PMID:10617198
 A;Accession: A85076
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-282 <STO>
 A;Cross-references: UNIPROT:Q9M0T7; GB:NC_001268; NID:G7267369; PIDN:CA577930.1; GSPDB:G
 C;Genetics:
 A;Gene: At4g07710
 A;Map position: 4

Query Match 40.3%; Score 50; DB 2; Length 282;
 Best Local Similarity 60.0%; Pred. No. 5.7; Mismatches 3; Indels 0; Gaps 0;
 Matches 9; Conservative 3;

QY 5 KPIADLYRGRESRPS 19
 |||||: : ||||
 DB 188 EPIADLFRKRRPS 202

RESULT 5

S28380
 A;Title: ribosomal protein L15.e - midge (Chironomus tentans)
 N;Alternate names: ribosomal protein Y110 homolog
 C;Species: Chironomus tentans
 C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
 C;Accession: S26380; S25633
 R;Galli, J.; Wieslander, L.
 Nucleic Acids Res. 20, 5473, 1992

A;Title: Structure of a gene in the dipteran Chironomus tentans encoding a yeast riboso
 A;Reference number: S26380; MUID:93065218; PMID:1437565

A;Accession: S26380
 A;Status: nucleic acid sequence not shown
 A;Molecule type: mRNA
 A;Residues: 1-204 <GAL>
 A;Cross-references: UNIPROT:P30736; EMBL:X68332; NID:g7116; PIDN:CAA48409.1; PID:g7117
 C;Superfamily: rat ribosomal protein L15
 C;Keywords: protein biosynthesis; ribosome

Query Match 39.1%; Score 48.5; DB 2; Length 204;
 Best Local Similarity 45.5%; Pred. No. 6.8; Mismatches 8; Indels 1; Gaps 1;
 Matches 10; Conservative 3;

QY 2 WQYKPIADLYRG-RESRPSAPR 22
 ||||: ||||: ||||
 DB 28 WQYQLTKLHRAPRSPDKAR 49

RESULT 6

T16141
 A;Title: Hypothetical protein F22F4.2 - Caenorhabditis elegans
 C;Species: Caenorhabditis elegans
 C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 02-Sep-2000
 C;Accession: T16141
 R;atireille, P.
 submitted to the EMBL Data Library, April 1996
 A;Description: The sequence of C. elegans cosmid F22F4.
 A;Reference number: Z18467
 A;Accession: T16141
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-432 <LAT>
 A;Cross-references: EMBL:U53343; NID:G1255881; PID:G1255883; PIDN:AAA96220.1; GSPDB:GNO
 A;Experimental source: strain Bristol N2; clone F22F4
 C;Genetics:
 A;Gene: CESP:F22F4.2
 A;Map position: X
 A;Introns: 8/2; 59/2; 386/3
 C;Superfamily: Passover protein

Query Match 38.7%; Score 48; DB 2; Length 432;
 Best Local Similarity 47.8%; Pred. No. 18; Mismatches 8; Indels 2; Gaps 1;
 Matches 11; Conservative 2;

QY 2 WQYKPIADLYRGRESRPSA--PR 22
 |||||: |||||: |||||
 DB 236 WGFHTFADLYAGRWQDSGVFPR 258

RESULT 7

T31343
 A;Title: Proline dehydrogenase (EC 1.5.99.8) - Bradyrhizobium japonicum
 C;Species: Bradyrhizobium japonicum
 C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
 C;Accession: T31343
 R;Straub, P.F.; Reynolds, P.H.; Althomsons, S.; Mett, V.; Zhu, Y.; Shearer, G.; Kohl, D.
 Appl. Environ. Microbiol. 62, 221-229, 1996
 A;Title: Isolation, DNA sequence analysis, and mutagenesis of a proline dehydrogenase g
 A;Reference number: Z21008; MUID:96141420; PMID:8572700
 A;Accession: T31343
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-1016 <STR>
 A;Cross-references: UNIPROT:Q59206; EMBL:X90401; NID:g1212778; PIDN:CAA62038.1; PID:g12

A;Gene: puta
 C;Superfamily: bifunctional proline dehydrogenase/delta-1-pyrroline-5-carboxylate dehyd
 C;Keywords: oxidoreductase

Query Match 37.9%; Score 47; DB 2; Length 1016;
 Best Local Similarity 50.0%; Pred. No. 64; Mismatches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;